Homogenization of locally stationary diffusions with possibly degenerate diffusion matrix

Rémi Rhodes
Université Paris-Dauphine, Ceremade
Place du Maréchal De Lattre De Tassigny
75775 Paris cedex 16 - FRANCE,
e-mail: rhodes@ceremade.dauphine.fr

February 10, 2009

Abstract : This paper deals with homogenization of second order divergence form parabolic operators with locally stationary coefficients. Roughly speaking, locally stationary coefficients have two evolution scales: both an almost constant microscopic one and a smoothly varying macroscopic one. The homogenization procedure aims to give a macroscopic approximation that takes into account the microscopic heterogeneities. This paper follows [13] and improves this latter work by considering possibly degenerate diffusion matrices.

Résumé : Nous étudions l'homogénéisation d'opérateurs paraboliques du second ordre sous forme divergence à coefficients localement stationnaires. Ces coefficients présentent deux échelles d'évolution: une évolution microscopique presque constante et une évolution macroscopique régulière. La théorie de l'homogénéisation consiste à donner une approximation macroscopique de l'opérateur initial qui tient compte des hétérogénéités microscopiques. Cet article fait suite à [13] et généralise ce dernier en considérant des matrices de diffusion pouvant dégénérer.

AMS classification: 60F17; (35B27; 35K65; 28D05).

1 Introduction

This paper follows [13] and deals with homogenization of second order PDEs with locally stationary coefficients by means of probabilistic tools. More precisely, we aim at describing the

asymptotic behavior, as ε goes to 0, of the following Stochastic Differential Equation (SDE)

$$(1) \qquad X_t^\varepsilon = x + \frac{1}{\varepsilon}\int_0^t b\big(\omega, \frac{X_r^\varepsilon}{\varepsilon}, X_r^\varepsilon\big)\, dr + \int_0^t c\big(\omega, \frac{X_r^\varepsilon}{\varepsilon}, X_r^\varepsilon\big)\, dr + \int_0^t \sigma\big(\omega, \frac{X_r^\varepsilon}{\varepsilon}, X_r^\varepsilon\big)\, dB_r,$$

where B is a standard d-dimensional Brownian motion and the parameter ω evolves in a random medium Ω, that is a probability space with suitable stationarity and ergodicity properties. For each fixed value of the parameter $y \in \mathbb{R}^d$, the coefficients $b(\omega, \cdot, y)$, $c(\omega, \cdot, y)$ and $\sigma(\omega, \cdot, y)$ are stationary random fields (the parameter ω stands for this randomness). That is why they are said to be locally stationary. The generator $\mathcal{L}^\varepsilon$ of the process X^ε can be written in divergence form as

$$(2) \qquad \mathcal{L}^\varepsilon = \frac{1}{2} e^{2V(x)} \sum_{i,j=1}^d \frac{\partial}{\partial x_i}\big(e^{-2V(x)}[a + H](\omega, \frac{x}{\varepsilon}, x)\frac{\partial}{\partial x_j}\,\big)$$

for an antisymmetric matrix H, a real-valued function V and $a = \sigma\sigma^*$.

Let us first briefly outline the chronological approach of this issue. The convergence of the previous SDE (or the connected PDE) has been first established in the locally periodic case, that is when the coefficients are deterministic and periodic with respect to the variable x/ε [1, 2]. Due to the lack of compactness of a random medium, the random case raises more difficulties. As far as we know, the first work in this context is due to Olla and Siri in [11]. The authors considered a nearest neighbors random walk on $\mathbb{Z}$ evolving in a locally stationary environment. They established an invariance principle for this process under diffusive scaling of space and time. The main tool of the proof is the explicit formula of the correctors, which only holds in dimension one under a strong diffusivity condition.

In [13], an alternative approach is suggested, which is not restricted to the dimension one. As in the locally periodic setting, the method is based on a local analysis of the microscopic behavior (corresponding to the variable x/ε) of the process X^ε to construct the so-called correctors and to identify the limiting process. However, unlike the locally periodic case, these correctors turn out to have bad asymptotic properties at a macroscopic scale, in the sense that the classical ergodic theory cannot describe their asymptotic behavior. Overcoming this issue is the main contribution of [13]. The main assumption is the uniform ellipticity of the matrix a, namely that there exits a constant $M > 0$ such that for all $x, y, X \in \mathbb{R}^d$,

$$\frac{1}{M}|X|^2 \leq (a(\omega, x, y)X, X) \leq M|X|^2.$$

This condition is very convenient for two reasons. From the dynamical angle, it ensures the local ergodicity of the process X^ε. From the technical angle, it provides strong estimates of the transition densities of the process X^ε as well as regularity properties of its generator. The control of the process X^ε, in particular its invariant measure and its tightness, is easily derived from this assumption.

In this present paper, we intend to improve this latter work by removing the uniform ellipticity assumption. It is replaced by microscopic ergodicity conditions (Assumption 2.5), which seem not too far from being minimal to apply classical ergodic theory and then pass to the limit in (1). The class of considered coefficients then includes possibly degenerate matrices a. In other words, we can treat diffusion coefficients a that may reduce to 0 along some directions. Under suitable assumptions, we will prove that the process X^ε converges to the solution $\bar{X}$ of a SDE with deterministic coefficients, whose generator can be rewritten in divergence form as

$$\bar{L} = (1/2)e^{2V(x)} \sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\big(e^{-2V(x)}[\bar{A}+\bar{H}](x)\frac{\partial}{\partial x_j}\big), \tag{3}$$

where the so-called homogenized coefficients $\bar{A}$ and $\bar{H}$ are respectively symmetric positive and antisymmetric. It is worth emphasizing that $\overline{A}$ may degenerate, even under strong non-degeneracy assumptions of the initial diffusion coefficient a. We will prove that the limiting diffusion is trapped in a fixed subspace of $\mathbb{R}^d$ and possesses strong diffusivity properties along this subspace.

We should finally point out that there are only a few papers dealing with possibly degenerate diffusion coefficients in the whole literature about probabilistic homogenization of SDEs. In the periodic setting, recent advances have been made by Hairer and Pardoux in [5]. Their approach deeply differs from ours. They allow the diffusion to be strongly degenerate in some area of the torus, and even to reduce to 0 over an open domain, provided that the diffusion quickly reaches a strongly regularizing area (typically, it satisfies a strong Hörmander type condition). Our approach does not allow locally such strong degeneracies but does not require any regularizing area. As a consequence, we can construct examples that are everywhere degenerate. Moreover, the technics used in [5] rely on the compactness of the torus and cannot be adapted to the random setting.

The structure of the paper is the following. In section 2, we introduce all the notations and assumptions. Our results are stated in Section 4 and an example is given in Section 5. The construction of the corrector is carried out in Section 6. Section 7 deals with the regularity properties of the process X^ε such as its invariant measure and the Itô formula. Section 8 is devoted to establishing the asymptotic properties of the process X^ε. Section 9 explains the proofs of the homogenization procedure. The tightness of the process X^ε is treated separately in Section 10.

2 Setup and Assumptions

Random medium. From now on, $d \geq 1$ is a fixed integer. Following [7], we introduce the following

Definition 2.1. *Let* $(\Omega, \mathcal{G}, \mu)$ *be a probability space and* $\left\{\tau_x; x \in \mathbb{R}^d\right\}$ *a group of measure preserving transformations acting ergodically on* Ω*:*

1) $\forall A \in \mathcal{G}, \forall x \in \mathbb{R}^d$*,* $\mu(\tau_x A) = \mu(A)$*,*

2) If for any $x \in \mathbb{R}^d$ $\tau_x A = A$*, then* $\mu(A) = 0$ *or* 1*,*

3) For any measurable function $\boldsymbol{g}$ *on* $(\Omega, \mathcal{G}, \mu)$*, the function* $(x, \omega) \mapsto \boldsymbol{g}(\tau_x \omega)$ *is measurable on* $(\mathbb{R}^d \times \Omega, \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{G})$*.*

The expectation with respect to the random medium is denoted by $\mathbb{M}$. Denote by $L^2(\Omega)$ the space of square integrable functions, by $|.|_2$ the corresponding norm and by $(.,.)_2$ the associated inner product. The operators defined on $L^2(\Omega)$ by $T_x\boldsymbol{f}(\omega) = \boldsymbol{f}(\tau_x\omega)$ form a strongly continuous group of unitary maps in $L^2(\Omega)$. For every function $\boldsymbol{f} \in L^2(\Omega)$, let $f(\omega, x) = \boldsymbol{f}(\tau_x\omega)$. Each function $\boldsymbol{f}$ in $L^2(\Omega)$ defines in this way a stationary ergodic random field on $\mathbb{R}^d$. In what follows we will use the bold type to denote an element $\boldsymbol{f} \in L^2(\Omega)$ and the normal type $f(\omega, x)$ (or even $f(x)$) to distinguish from the associated stationary field. The group possesses d generators (throughout this paper, e_i stands for the i-th vector of the canonical basis of $\mathbb{R}^d$)

$$D_i\boldsymbol{g} = \lim_{h\to 0} \frac{T_{he_i}\boldsymbol{g} - \boldsymbol{g}}{h} \text{ if exists,} \tag{4}$$

which are closed and densely defined. Setting

$$\mathcal{C} = \operatorname{Span}\left\{\boldsymbol{g} \star \varphi; \boldsymbol{g} \in L^\infty(\Omega), \varphi \in C_c^\infty(\mathbb{R}^d)\right\}, \text{ with } \boldsymbol{g} \star \varphi(\omega) = \int_{\mathbb{R}^d} \boldsymbol{g}(\tau_x\omega)\varphi(x)\,dx, \tag{5}$$

the space $\mathcal{C}$ is dense in $L^2(\Omega)$ and $\mathcal{C} \subset \operatorname{Dom}(D_i)$ for all $1 \leq i \leq d$, with $D_i(\boldsymbol{g} \star \varphi) = -\boldsymbol{g}\star\partial\varphi/\partial x_i$. If $\boldsymbol{g} \in \operatorname{Dom}(D_i)$, we also have $D_i(\boldsymbol{g}\star\varphi) = D_i\boldsymbol{g}\star\varphi$. For $\boldsymbol{f} \in \bigcap_{i=1}^d \operatorname{Dom}(D_i)$, we define the divergence operator Div by $\operatorname{Div}\boldsymbol{f} = \sum_{i=1}^d D_i\boldsymbol{f}$. We distinguish this latter operator from the usual divergence operator on $\mathbb{R}^d$ denoted by the small type div.

Locally stationary random fields. Following the notations introduced just above, for a measurable function $\boldsymbol{f} : \Omega \times \mathbb{R}^d \to \mathbb{R}^n$, ($n \geq 1$), we can consider the associated locally stationary random field $(x, y) \mapsto \mathbf{f}(\tau_x\omega, y) = f(\omega, x, y)$ (or even $f(x, y)$).

Structure of the coefficients. The coefficients $\boldsymbol{\sigma} : \Omega \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$, $\boldsymbol{H} : \Omega \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$, $\tilde{\boldsymbol{\sigma}} : \Omega \to \mathbb{R}^{d\times d}$ and $V : \mathbb{R}^d \to \mathbb{R}$ denote measurable functions with respect to the underlying product σ-fields. As explained above, $\boldsymbol{\sigma}$ and $\boldsymbol{H}$ define locally stationary random fields and $\tilde{\boldsymbol{\sigma}}$ a stationary random field. $\boldsymbol{H}$ is antisymmetric. We define two new matrix-valued functions by $\boldsymbol{a} = \boldsymbol{\sigma}\boldsymbol{\sigma}^*$ and $\tilde{\boldsymbol{a}} = \tilde{\boldsymbol{\sigma}}\tilde{\boldsymbol{\sigma}}^*$. Furthermore, for some positive constant Λ, the coefficients $\boldsymbol{\sigma}$, $\boldsymbol{H}$, $\tilde{\boldsymbol{\sigma}}$ and V satisfy

Assumption 2.2. (Regularity). *For each fixed* $\omega \in \Omega$*, the coefficients* $\sigma(\omega, ., .)$*,* $H(\omega, ., .)$ *and* $\tilde{\sigma}(\omega, .)$ *are two times continuously differentiable with respect to each variable and are, as well as their derivatives up to order two,* Λ*-Lipschitzian and bounded by* Λ*.* V *is three times*

continuously differentiable and is, as well as its derivatives up to order three, bounded by Λ and Λ-Lipschitzian.

Let us now describe the degeneracies of the matrix $\boldsymbol{a}$. Roughly speaking, the degeneracies of $\boldsymbol{a}$ are assumed to be controlled by the reference matrix $\tilde{\boldsymbol{a}}$. To be more explicit, let us first introduce the

Definition 2.3. *Given a $d \times d$ matrix-valued function $g : \mathbb{R}^d \to \mathbb{R}^{d\times d}$, a $d \times d$ symmetric matrix A and a real $C > 0$, g is said to be (C, A)-controlled if $\forall y, y' \in \mathbb{R}^d$*

$$|g(y)| \leq CA, \quad \text{and } |g(y) - g(y')| \leq CA|y - y'|,$$

where $|M| = (MM^*)^{1/2}$ stands for the absolute value of the matrix M (given 2 symmetric matrices A, B, the relation $A \leq B$ means that the matrix $B - A$ is symmetric positive).

We now precise the control of $\boldsymbol{a}$ by $\tilde{\boldsymbol{a}}$:

Assumption 2.4. (Control). *We assume that*

$$M^{-1}\widetilde{\boldsymbol{a}}(\omega) \leq \boldsymbol{a}(\omega, y) \leq M\widetilde{\boldsymbol{a}}(\omega)$$

for some strictly positive constant M and for every $(\omega, y) \in \Omega \times \mathbb{R}^d$. Moreover, for any $i, j \in \{1, \dots, d\}$ and $(\omega, y) \in \Omega \times \mathbb{R}^d$, the matrices $\partial_{y_i}\boldsymbol{a}(\omega, y)$, $\partial^2_{y_iy_j}\boldsymbol{a}(\omega, y)$, $\boldsymbol{H}(\omega, y)$, $\partial_{y_i}\boldsymbol{H}(\omega, y)$, $\partial^2_{y_iy_j}\boldsymbol{H}(\omega, y)$ are $(M, \tilde{\boldsymbol{a}}(\omega))$-controlled. We further assume that

$$|\boldsymbol{\sigma}(\omega, y + h) - \boldsymbol{\sigma}(\omega, y))|^2 \leq M\widetilde{\boldsymbol{a}}(\omega)|h|^2$$

for any $y, h \in \mathbb{R}^d$ and that $\int_{\mathbb{R}^d} e^{-2V(y)}\, dy = 1$.

To ensure the local ergodicity of the process X^ε, we make the following assumption:

Assumption 2.5 (Ergodicity). *Let us consider the Friedrich extension (see [4, p. 53] or Section 5) of the symmetric operator $\tilde{\boldsymbol{S}}$ defined on $\mathcal{C} \subset L^2(\Omega)$ by $\tilde{\boldsymbol{S}} = (1/2)\sum_{i,j=1}^d D_i(\tilde{\boldsymbol{a}}_{i,j}D_j)$. This extension, still denoted $\tilde{\boldsymbol{S}}$, is self-adjoint. We then assume that the semi-group generated by $\tilde{\boldsymbol{S}}$ is ergodic, that is its invariant functions are μ almost surely constant (see e.g. Rhodes [12]).*

Remark. *Assumptions 2.2 may appear restrictive and can surely be relaxed (see [3] for results in this direction in the context of quasilinear PDEs). In particular, the statement of the homogenization property only involves the derivatives of order one with respect to $y \in \mathbb{R}^d$ (see Theorem 3.1). However, it avoids dealing with heavy regularizing procedures that are not the purpose of this work.*

Diffusion in a locally ergodic environment. For $j = 1, \dots, d$, we define the coefficients

(6)
$$\boldsymbol{b}_j(\omega, y) = \frac{1}{2}\sum_{i=1}^d D_i(\boldsymbol{a} + \boldsymbol{H})_{ij}(\omega, y), \quad \boldsymbol{c}_j(\omega, y) = \frac{e^{2V(y)}}{2}\sum_{i=1}^d \partial_{y_i}\big(e^{-2V}[\boldsymbol{a} + \boldsymbol{H}]_{ij}\big)(\omega, y).$$

From Assumption 2.2, the functions $b_j(\omega,.,.)$ and $c_j(\omega,.,.)$ are Lipschitzian so that, for a starting point $x \in \mathbb{R}^d$ and $\varepsilon > 0$, we can consider the strong solution X^ε of the following Stochastic Differential Equation (SDE) with locally stationary coefficients:

$$X_t^\varepsilon = x + \frac{1}{\varepsilon}\int_0^t b\left(\overline{X}_r^\varepsilon, X_r^\varepsilon\right)\, dr + \int_0^t c\left(\overline{X}_r^\varepsilon, X_r^\varepsilon\right)\, dr + \int_0^t \sigma\left(\overline{X}_r^\varepsilon, X_r^\varepsilon\right)\, dB_r, \tag{7}$$

where we have set $\overline{X}_t^\varepsilon \equiv X_t^\varepsilon/\varepsilon$ and B is a standard d-dimensional Brownian motion (the random medium and the Brownian motion are independent). We point out that the generator of this diffusion could be formally written in divergence form as

$$\mathcal{L}^\varepsilon = \frac{1}{2}e^{2V(x)}\sum_{i,j=1}^d \frac{\partial}{\partial x_i}\big(e^{-2V(x)}[a+H](\omega, x/\varepsilon, x)\frac{\partial}{\partial x_j}\,\big). \tag{8}$$

Notations. *For the sake of simplicity, we indicate the starting point x of X^ε by writing, when necessary, $\mathbb{P}_x^\varepsilon$ (and $\mathbb{E}_x^\varepsilon$ for the corresponding expectation), this avoids heavy notations as $X^{\varepsilon,x}$. We can then consider the probability measure $\bar{\mathbb{P}}^\varepsilon \equiv \mathbb{M}\int_{\mathbb{R}^d}\mathbb{P}_x^\varepsilon[.]e^{-2V(x)}\,dx$ and its expectation $\bar{\mathbb{E}}^\varepsilon$.*

3 Main Results

Let us now state the main result of this paper. Under the previous assumptions, we can prove

Theorem 3.1. Homogenization. *The law $\bar{\mathbb{P}}^\varepsilon$ of the process X^ε weakly converges in $C([0,T];\mathbb{R}^d)$ towards the law of the process X that solves the following SDE with deterministic coefficients (they do not depend on the medium Ω):*

$$X_t = x + \int_0^t \overline{B}(X_r)\,dr + \int_0^t \overline{A}^{1/2}(X_r)\,dB_r. \tag{9}$$

The coefficients $\overline{A}$ and $\overline{B}$ are of class C^2 and are defined, for $y \in \mathbb{R}^d$, by

$$\overline{A}(y) = \lim_{\lambda\to 0}\mathbb{M}[(I+\boldsymbol{D}\boldsymbol{u}_\lambda)^*\boldsymbol{a}(I+\boldsymbol{D}\boldsymbol{u}_\lambda)(.,y)], \tag{10a}$$

$$\overline{H}(y) = \lim_{\lambda\to 0}\mathbb{M}[(I+\boldsymbol{D}\boldsymbol{u}_\lambda)^*\boldsymbol{H}(I+\boldsymbol{D}\boldsymbol{u}_\lambda)(.,y)], \tag{10b}$$

$$\overline{B}(y) = (1/2)e^{2V(y)}\partial_y(e^{-2V}[\overline{A}+\overline{H}])(y). \tag{10c}$$

Formally speaking, for each $y \in \mathbb{R}^d$ and $\lambda > 0$, the entries $\left(\boldsymbol{u}_\lambda^i(.,y)\right)_{1\le i\le d}$ of the function $\boldsymbol{u}_\lambda(.,y) : \Omega \to \mathbb{R}^d$ solve the following so-called auxiliary problems, which are stated on the random medium

$$\lambda\boldsymbol{u}_\lambda^i(.,y) - \frac{1}{2}\sum_{j,k} D_j\left[(\boldsymbol{a}_{jk}+\boldsymbol{H}_{jk})D_k\boldsymbol{u}_\lambda^i(.,y)\right] = \boldsymbol{b}_i(.,y).$$

__Remark__. A rigorous description of $\boldsymbol{u}_\lambda(.,y)$ is given in Section 6. In particular, in this degenerate framework, the "gradients" $\boldsymbol{D}\boldsymbol{u}_\lambda$ do not exist but along the direction $\tilde{\boldsymbol{\sigma}}$, that is the only expression $\tilde{\boldsymbol{\sigma}}^*\boldsymbol{D}\boldsymbol{u}_\lambda$ can be given a rigorous sense. Because of the control of $\boldsymbol{a}$ and $\boldsymbol{H}$ by $\tilde{\boldsymbol{a}}$ (Assumption 2.4), it then makes sense to consider formulae (10a) *and* (10b) *(see Section 6 for further details).*

Since the diffusion coefficient $\boldsymbol{a}$ is allowed to degenerate, the reader may wonder whether the homogenized diffusion coefficient may also degenerate. The following proposition details the structure of the limiting diffusion coefficient $\bar{A}$:

Proposition 3.2. Geometry of the homogenized coefficients. *The kernel $K = \mathrm{Ker}(\bar{A}(y))$ of $\bar{A}(y)$ does not depend on the point $y \in \mathbb{R}^d$ where it is computed. For each $y \in \mathbb{R}^d$, $\overline{B}(y) \in K^\perp$ ($K^\perp$ is the orthogonal complement to K) and there exists a constant $\alpha_{3.2} > 0$, such that*

$$\forall y \in \mathbb{R}^d,\ \forall x \in K^\perp,\ \alpha_{3.2}^{-1}|x|^2\langle x, \bar{A}(y)x\rangle \leq \alpha_{3.2}|x|^2.$$

In other words, for each starting point $x \in \mathbb{R}^d$, the limiting process X (see (9)*) can be seen as the solution of a SDE defined on $x + K^\perp$ with a uniformly elliptic diffusion matrix $\bar{A}$.*

4 Example

Let us consider a simple example in the two dimensional 2π-periodic case. The 2-dimensional torus $\mathbb{T}^2$ is seen as the random medium equipped with the induced Lebesgue measure, still denoted by μ to stick with the notations of the paper . We aim at constructing a degenerate homogenized coefficient. For this purpose, let us first define

$$\forall x \in \mathbb{R}^2, \quad \widetilde{\boldsymbol{\sigma}}(x) = \begin{pmatrix} 1 & 1/c \\ c & 1 \end{pmatrix},$$

where $c \notin \pi\mathbb{Q}$ is a constant, and $\widetilde{\boldsymbol{a}} = \widetilde{\boldsymbol{\sigma}}\widetilde{\boldsymbol{\sigma}}^*$. Choose now any smooth function $\boldsymbol{U} : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$, with bounded derivatives up to order 2, 2π-periodic with respect to its first argument $x \in \mathbb{R}^2$ and satisfying

$$\forall (x,y) \in \mathbb{R}^2 \times \mathbb{R}^2, \quad M^{-1}\mathrm{Id} \leq \boldsymbol{U}\boldsymbol{U}^*(x,y) \leq M\,\mathrm{Id}.$$

Define $\forall (x,y) \in \mathbb{R}^2 \times \mathbb{R}^2$, $V(y) = e^{-|y|^2}/\pi$, $\boldsymbol{\sigma}(x,y) = \widetilde{\boldsymbol{\sigma}}\boldsymbol{U}(x,y)$ and $\boldsymbol{H} = 0$. Let us check that these coefficients satisfy all our assumptions. From the smoothness of the coefficients, it is plain to see that Assumptions 2.4 and 2.2 are fulfilled. Assumption 2.5 results from the Weyl equipartition theorem ($c \notin \pi\mathbb{Q}$). Theorem 3.1 thus holds.

Let us now prove that $\bar{A}$ is degenerate and does not trivially reduce to 0. Let us denote by $\widetilde{A}$ the homogenized coefficient associated to $\tilde{\boldsymbol{a}}$. From the proof of Proposition 3.2, for any $y \in \mathbb{R}^2$ and $X \in \mathbb{R}^2$, we have

$$C^{-1}\big\langle X, \widetilde{A}X\big\rangle \leq \big\langle X, \bar{A}(y)X\big\rangle \leq C\big\langle X, \widetilde{A}X\big\rangle = 0.$$

So we just have to compute $\widetilde{A}$. Since $\widetilde{\boldsymbol{\sigma}}$ is constant, it is straightforward to check that $\widetilde{A}$ actually matches $\widetilde{\boldsymbol{\sigma}}\widetilde{\boldsymbol{\sigma}}^*$ with the help of (45). Indeed, for a given smooth function $\boldsymbol{\varphi}$ defined on $\mathbb{T}^2$ and $x \in \mathbb{R}^2$, the right-hand side of (45) expands as

$$\begin{aligned}\mathbb{M}[|\widetilde{\boldsymbol{\sigma}}(D\boldsymbol{\varphi}+x)|^2] =&\mathbb{M}[|\widetilde{\boldsymbol{\sigma}}^* D\boldsymbol{\varphi}|^2] + 2\left\langle \widetilde{\boldsymbol{\sigma}}^* x, \widetilde{\boldsymbol{\sigma}}^*\mathbb{M}[D\boldsymbol{\varphi}]\right\rangle + \left\langle \widetilde{\boldsymbol{\sigma}}^* x, \widetilde{\boldsymbol{\sigma}}^* x\right\rangle \\ =&\mathbb{M}[|\widetilde{\boldsymbol{\sigma}}^* D\boldsymbol{\varphi}|^2] + \left\langle \widetilde{\boldsymbol{\sigma}}^* x, \widetilde{\boldsymbol{\sigma}}^* x\right\rangle .\end{aligned}$$

The infimum is then clearly reached for $\boldsymbol{\varphi} = 0$.

Finally, we let the reader check that $\widetilde{A} = \widetilde{\boldsymbol{\sigma}}\widetilde{\boldsymbol{\sigma}}^*$ does not reduce to 0 and that the vector $X_K = [1 \quad -c]^*$ satisfies $\widetilde{A}X_K = 0$.

In a general way, because of the various geometries of random media, it is not clear whether $\bar{A}$ is degenerate or not. The reader may find in [3] examples (in a slightly different framework) where the diffusion matrix reduces to 0 though the diffusion coefficient $\boldsymbol{\sigma}$ is elliptic over a set of full Lebesgue measure, and conversely, an example where $\boldsymbol{\sigma}$ degenerates and $\bar{A}$ is uniformly elliptic. $\square$

5 Construction of unbounded operators

Throughout this paper, we will need to construct suitable extensions of unbounded operators defined on a dense subspace of a given L^2-space. This construction is always the same and follows [4, Ch. 3, Sect 3.] or [9, Ch. 1, Sect 2.], to which the reader is referred for further details than those given below. That is the reason why we explain it in a generic way. We also point out that the Friedrich extension of $\tilde{\boldsymbol{S}}$ (see Assumption 2.5) corresponds to this construction.

Consider a probability space Ω equipped with a probability measure $\mathbb{P}$, a dense subspace $\mathcal{D}$ of $L^2(\Omega;\mathbb{P})$, a positive symmetric bilinear form $\langle\cdot,\cdot\rangle$ defined on $\mathcal{D}\times\mathcal{D}$ ($\|\cdot\|$ denotes the corresponding semi-norm) and a bilinear form B on $\mathcal{D}\times\mathcal{D}$ that satisfies for any $\varphi,\psi\in\mathcal{D}$

$$\alpha^{-1}\|\varphi\|^2 \le B(\varphi,\varphi), \quad B(\varphi,\psi)\le \alpha\|\varphi\|\|\psi\| \tag{11}$$

for some positive constant $\alpha > 0$. Let us denote $(\cdot,\cdot)_2$ the canonical inner product on $L^2(\Omega;\mathbb{P})$.

From now on, we will say that the unbounded operator L on $L^2(\Omega;\mathbb{P})$ is constructed from $(\Omega,\mathbb{P},\langle\cdot,\cdot\rangle, B)$ if it is constructed as follows. We consider the inner product Π on $\mathcal{D}\times\mathcal{D}$ defined by

$$\Pi(\varphi,\psi) = (\varphi,\psi)_2 + \langle\varphi,\psi\rangle$$

and the closure $\mathbb{H}$ of $\mathcal{D}$ with respect to the corresponding norm. For each $\lambda > 0$, the bilinear form B_λ is defined on $\mathcal{D}\times\mathcal{D}$ by

$$B_\lambda(\varphi,\psi) = \lambda(\varphi,\psi)_2 + B(\varphi,\psi).$$

From (11), B_λ obviously extends to $\mathbb{H}\times\mathbb{H}$ (this extension is still denoted by B_λ). Furthermore, it is continuous and coercive on $\mathbb{H}\times\mathbb{H}$. Thus it defines a resolvent operator $G_\lambda : L^2(\Omega,\mathbb{P}) \to \mathbb{H}$,

which is one-to-one. We can then define L as $\lambda - G_\lambda^{-1}$ with domain $\mathrm{Dom}(L) = G_\lambda(L^2(\Omega, \mathbb{P}))$. This definition does not depend on $\lambda > 0$. It is readily seen that a function $\varphi \in \mathbb{H}$ belongs to $\mathrm{Dom}(L)$ if and only if the map $\psi \in \mathbb{H} \mapsto B_\lambda(\varphi, \psi)$ is $L^2(\Omega, \mathbb{P})$ continuous. In this case, we can find $f \in L^2(\Omega, \mathbb{P})$ such that $B_\lambda(\varphi, \cdot) = (f, \cdot)_2$. Then $L\varphi$ exactly matches $f - \lambda\varphi$. Note that $B(\varphi, \psi) = -(L\varphi, \psi)_2$ for any $\varphi \in \mathrm{Dom}(L)$ and $\psi \in \mathbb{H}$. We point out that the unbounded operator L is closed and densely defined. Moreover, its adjoint operator L^* in $L^2(\Omega; \mathbb{P})$ coincides with the operator constructed from $(\Omega, \mathbb{P}, \langle \cdot, \cdot \rangle, \check{B})$, where the bilinear form $\check{B}$ is defined on $\mathcal{D} \times \mathcal{D}$ by $\check{B}(\varphi, \psi) = B(\psi, \varphi)$. As a consequence $(L^*)^* = L$.

Notations. *In what follows, the notation* $(\mathbb{H}, L, \mathrm{Dom}(L), (G_\lambda)_{\lambda>0}) = \Xi((\Omega, \mathbb{P}, \langle \cdot, \cdot \rangle, B))$ *means that* $\mathbb{H}$, L, $\mathrm{Dom}(L)$, $(G_\lambda)_{\lambda>0}$ *are constructed from* $(\Omega, \mathbb{P}, \langle \cdot, \cdot \rangle, B)$ *as explained above.*

6 Auxiliary Problems

Setup and notations. Let us now focus on the different operators induced on the random medium Ω by the matrices $\boldsymbol{a}(\cdot, y)$ and $\boldsymbol{H}(\cdot, y)$, for each $y \in \mathbb{R}^d$. We aim at extending the following operators defined on $\mathcal{C}$ by

$$\boldsymbol{S}^y \equiv \frac{1}{2}\sum_{i,j=1}^{d} D_i\big(\boldsymbol{a}_{ij}(\cdot, y)D_j\big), \quad \boldsymbol{L}^y \equiv \frac{1}{2}\sum_{i,j=1}^{d} D_i\big((\boldsymbol{a}+\boldsymbol{H})_{ij}(., y)D_j\big), \tag{12}$$

according to the method detailed in Section 5.

The positive symmetric bilinear form $(\cdot, \cdot)_1$ is defined on $\mathcal{C} \times \mathcal{C}$ by

$$(\boldsymbol{\varphi}, \boldsymbol{\psi})_1 \equiv -(\boldsymbol{\varphi}, \tilde{\boldsymbol{S}}\boldsymbol{\psi})_2 = (1/2)\big(\tilde{\boldsymbol{a}} D\boldsymbol{\varphi}, D\boldsymbol{\psi}\big)_2, \tag{13}$$

and the associated seminorm $\|\cdot\|_1$ by $\|\boldsymbol{\varphi}\|_1^2 \equiv (\boldsymbol{\varphi}, \boldsymbol{\varphi})_1$.

For any $\boldsymbol{\varphi}, \boldsymbol{\psi} \in \mathcal{C}$, we define the bilinear forms (y is fixed)

$$\mathcal{B}^S(\boldsymbol{\varphi}, \boldsymbol{\psi}) \equiv -(\boldsymbol{S}^y\boldsymbol{\varphi}, \boldsymbol{\psi})_2 = (1/2)(\boldsymbol{a}(\cdot, y)D\boldsymbol{\varphi}, D\boldsymbol{\psi})_2,$$
$$\mathcal{B}^L(\boldsymbol{\varphi}, \boldsymbol{\psi}) \equiv -(\boldsymbol{L}^y\boldsymbol{\varphi}, \boldsymbol{\psi})_2 = (1/2)((\boldsymbol{a}+\boldsymbol{H})(\cdot, y)D\boldsymbol{\varphi}, D\boldsymbol{\psi})_2.$$

From Assumption 2.4 and the antisymmetry of $\boldsymbol{H}$, it is readily seen that $M^{-1}\|\boldsymbol{\varphi}\|_1^2 \le \mathcal{B}^S(\boldsymbol{\varphi}, \boldsymbol{\varphi})$ (resp. $M^{-1}\|\boldsymbol{\varphi}\|_1^2 \le \mathcal{B}^L(\boldsymbol{\varphi}, \boldsymbol{\varphi})$) and $\mathcal{B}^S(\boldsymbol{\varphi}, \boldsymbol{\psi}) \le M\|\boldsymbol{\varphi}\|_1\|\boldsymbol{\psi}\|_1$ (resp. $\mathcal{B}^L(\boldsymbol{\varphi}, \boldsymbol{\psi}) \le 2M\|\boldsymbol{\varphi}\|_1\|\boldsymbol{\psi}\|_1$). We can then define

$$(\mathbb{H}_1, \boldsymbol{S}^y, \mathrm{Dom}(\boldsymbol{S}^y), (G_\lambda^{S^y})_{\lambda>0}) = \Xi(\Omega, \mu, (\cdot, \cdot)_1, \mathcal{B}^S),$$
$$(\mathbb{H}_1, \boldsymbol{L}^y, \mathrm{Dom}(\boldsymbol{L}^y), (G_\lambda^{L^y})_{\lambda>0}) = \Xi(\Omega, \mu, (\cdot, \cdot)_1, \mathcal{B}^L).$$

Let us additionally denote by $(\boldsymbol{L}^y)^*$ the adjoint operator of $\boldsymbol{L}^y$ in $L^2(\Omega)$. Note that $\boldsymbol{S}^y$ is self-adjoint.

We define the space $\mathbb{D}$ as the closure in $(L^2(\Omega))^d$ of the set $\{\tilde{\boldsymbol{\sigma}}^* D\boldsymbol{\varphi}; \boldsymbol{\varphi} \in \mathcal{C}\}$. We point out that, whenever $\boldsymbol{\varphi}, \boldsymbol{\psi}$ belong to $\mathcal{C}$, $2(\boldsymbol{\varphi}, \boldsymbol{\psi})_1 = (\tilde{\boldsymbol{\sigma}}^* D\boldsymbol{\varphi}, \tilde{\boldsymbol{\sigma}}^* D\boldsymbol{\psi})_2$, so that the application $\Theta : \mathcal{C} \to \mathbb{D}$, $\boldsymbol{\varphi} \mapsto \tilde{\boldsymbol{\sigma}}^* D\boldsymbol{\varphi}$ can be extended to the whole space $\mathbb{H}_1$. For each function $\boldsymbol{f} \in \mathbb{H}_1$, we will note $\nabla^{\tilde{\sigma}} \boldsymbol{f}$ for $\Theta(\boldsymbol{f})$ and this represents in a way the gradient of the function $\boldsymbol{f}$ along the direction $\tilde{\boldsymbol{\sigma}}$. Similarly, for each fixed $y \in \mathbb{R}^d$, we define for any $\boldsymbol{\varphi} \in \mathbb{H}_1$ the gradient along the direction $\boldsymbol{\sigma}(\cdot, y)$. It will be denoted by $\nabla^{\sigma(\cdot,y)}\boldsymbol{\varphi}$ and is equal to $\boldsymbol{\sigma}(\cdot, y)^* D\boldsymbol{\varphi}$ for any $\boldsymbol{\varphi} \in \mathcal{C}$. From Assumption 2.4, for each $\boldsymbol{\varphi} \in \mathbb{H}_1$, the mapping $y \in \mathbb{R}^d \mapsto \nabla^{\sigma(\cdot,y)}\boldsymbol{\varphi} \in \mathbb{D}$ is continuous:

$$\forall (y, h) \in (\mathbb{R}^d)^2, \quad |\nabla^{\sigma(\cdot,y+h)}\boldsymbol{\varphi} - \nabla^{\sigma(\cdot,y)}\boldsymbol{\varphi}|_2^2 \le M|h|^2 \|\boldsymbol{\varphi}\|_1^2. \tag{14}$$

For $y \in \mathbb{R}^d$ and $\boldsymbol{\varphi}, \boldsymbol{\psi} \in \mathcal{C}$, we derive from Assumption 2.4

$$(\boldsymbol{L}^y \boldsymbol{\varphi}, \boldsymbol{\psi})_2 = -\frac{1}{2}\big(D\boldsymbol{\varphi}, (\boldsymbol{a} + \boldsymbol{H})(\cdot, y) D\boldsymbol{\psi}\big)_2 \le C|\nabla^{\tilde{\sigma}}\boldsymbol{\varphi}|_2 |\nabla^{\tilde{\sigma}}\boldsymbol{\psi}|_2, \tag{15}$$

so that we can define a bilinear form $\boldsymbol{T}^y$ on the whole space $\mathbb{D} \times \mathbb{D}$ such that $\forall \boldsymbol{\varphi}, \boldsymbol{\psi} \in \mathcal{C}$

$$-(\boldsymbol{L}^y \boldsymbol{\varphi}, \boldsymbol{\psi})_2 = \boldsymbol{T}^y(\nabla^{\tilde{\sigma}}\boldsymbol{\varphi}, \nabla^{\tilde{\sigma}}\boldsymbol{\psi}). \tag{16}$$

Thanks to Assumption 2.2, we can consider the differential $\partial \boldsymbol{T}^y$ of $\boldsymbol{T}^y$ defined, for $\boldsymbol{\varphi}, \boldsymbol{\psi} \in \mathcal{C}$, by $\partial \boldsymbol{T}^y(\boldsymbol{\varphi}, \boldsymbol{\psi}) = \partial_y(\boldsymbol{T}^y(\boldsymbol{\varphi}, \boldsymbol{\psi}))$. From Assumption 2.4 and similarly to (15), $\partial \boldsymbol{T}^y$ extends to $\mathbb{D} \times \mathbb{D}$. From Assumption 2.4, it is then plain to see that the relation $\partial_y\big(\boldsymbol{T}^y(\boldsymbol{\xi}, \boldsymbol{\zeta})\big) = \partial \boldsymbol{T}^y(\boldsymbol{\xi}, \boldsymbol{\zeta})$ still holds for $\boldsymbol{\xi}, \boldsymbol{\zeta} \in \mathbb{D}$.

Whenever a function $\boldsymbol{b}$ satisfies the property:

$$\exists C > 0, \forall \boldsymbol{\varphi} \in \mathcal{C}, \quad (\boldsymbol{b}, \boldsymbol{\varphi})_2 \le C\|\boldsymbol{\varphi}\|_1, \tag{17}$$

we will say that $\boldsymbol{b} \in \mathbb{H}_{-1}$ and we will define $\|\boldsymbol{b}\|_{-1}$ as the smallest constant C satisfying this property.

Solvability and regularity of the resolvent equation. For $\boldsymbol{h} \in L^2(\Omega)$, $\boldsymbol{u}_\lambda(., y) \equiv G_\lambda^{L^y}\boldsymbol{h}$ belongs to $\mathbb{H}_1 \cap \mathrm{Dom}(\boldsymbol{L}^y)$ and satisfies $\lambda \boldsymbol{u}_\lambda(\cdot, y) - \boldsymbol{L}^y \boldsymbol{u}_\lambda(\cdot, y) = \boldsymbol{h}$. Suppose that the right-hand side $\boldsymbol{h} = \boldsymbol{h}(\cdot, y)$ depends on the parameter $y \in \mathbb{R}^d$. We now investigate the y-regularity of $\boldsymbol{u}^\lambda(\cdot, y)$ from the regularity of $y \mapsto \boldsymbol{h}(\cdot, y)$ with respect to the norms $|\cdot|_2$ and $\|\cdot\|_{-1}$. We claim

Proposition 6.1. *Let us consider $\boldsymbol{h} : y \in \mathbb{R}^d \mapsto \boldsymbol{h}(., y) \in L^2(\Omega)$ and $\boldsymbol{f} : y \in \mathbb{R}^d \mapsto \boldsymbol{f}(., y) \in L^2(\Omega) \cap \mathbb{H}_{-1}$. Suppose that there exist C_2, C_{-1} such that:*

1) the application $y \mapsto \boldsymbol{h}(., y) \in L^2(\Omega)$ is two times continuously differentiable in $L^2(\Omega)$. The derivatives up to order 2 are bounded by C_2 in $L^2(\Omega)$ and are C_2-Lipschitz in $L^2(\Omega)$.

2) the application $y \mapsto \boldsymbol{f}(., y) \in L^2(\Omega) \cap \mathbb{H}_{-1}$ is two times continuously differentiable in $\mathbb{H}_{-1}$. The derivatives up to order 2 are bounded by C_{-1} in $\mathbb{H}_{-1}$ and are C_{-1}-Lipschitz in $\mathbb{H}_{-1}$.

Then, for any $\lambda > 0$, the solution $\boldsymbol{u}_\lambda(., y) \in \mathbb{H}_1 \cap \mathrm{Dom}(\boldsymbol{L}^y)$ of the equation

$$\lambda \boldsymbol{u}_\lambda(., y) - \boldsymbol{L}^y \boldsymbol{u}_\lambda(., y) = \boldsymbol{h}(., y) + \boldsymbol{f}(., y) \tag{18}$$

is two times continuously differentiable in $\mathbb{H}_1$ with respect to the parameter $y \in \mathbb{R}^d$. Furthermore there exists a constant $D_{6.1} > 0$, which only depends on M, C_{-1}, such that the functions $\boldsymbol{g}_\lambda(.,y) = \boldsymbol{u}_\lambda(.,y)$, $\partial_y \boldsymbol{u}_\lambda(.,y)$, $\partial^2_{yy}\boldsymbol{u}_\lambda(.,y)$ satisfy the property: $\forall (y,h) \in \mathbb{R}^2$,

$$\lambda|\boldsymbol{g}_\lambda(.,y)|_2^2 + \|\boldsymbol{g}_\lambda(.,y)\|_1^2 \leq D_{6.1}(1 + C_2^2/\lambda), \tag{19a}$$

$$\lambda|\boldsymbol{g}_\lambda(.,y+h) - \boldsymbol{g}_\lambda(.,y)|_2^2 + \|\boldsymbol{g}_\lambda(.,y+h) - \boldsymbol{g}_\lambda(.,y)\|_1^2 \leq D_{6.1}(1 + C_2^2/\lambda)|h|^2. \tag{19b}$$

Proof: The proof is readily adapted from [13, Prop. 4.1]. The method consists in differentiating the resolvent equation (18) with respect to the parameter $y \in \mathbb{R}^d$. In the uniformly elliptic setup [13, Prop. 4.1], this can be carried out thanks to the differentiability and the boundedness of $\boldsymbol{a}, \boldsymbol{H}$ and their derivatives up to order 2. In the degenerate setup, we need to control the matrices $\boldsymbol{a}$ and $\boldsymbol{H}$, as well as their derivatives up to order 2 with respect to the parameter $y \in \mathbb{R}^d$, by the matrix $\widetilde{\boldsymbol{a}}$ (see Assumption 2.4) in order to differentiate the function $y \mapsto \boldsymbol{u}_\lambda(\cdot, y)$ in $\mathbb{H}_1$. $\square$

Auxiliary problems: construction of the correctors. The end of this section is now devoted to the study of the solutions of the so-called auxiliary problems, that means the solutions $\boldsymbol{u}^i_\lambda(.,y)$ $(i = 1, \ldots, d)$ of the resolvent equations

$$\lambda \boldsymbol{u}^i_\lambda(.,y) - \boldsymbol{L}^y \boldsymbol{u}^i_\lambda(.,y) = \boldsymbol{b}_i(.,y), \tag{20}$$

where $\boldsymbol{b}_i(.,y) = (1/2)\sum_{j=1}^d D_j\big[(\boldsymbol{a} + \boldsymbol{H})_{ji}(.,y)\big]$. The weak form of the resolvent equation then reads for $\boldsymbol{\varphi} \in \mathcal{C}$

$$\lambda(\boldsymbol{u}^i_\lambda(.,y), \boldsymbol{\varphi})_2 + \boldsymbol{T}^y(\nabla^{\widetilde{\sigma}} \boldsymbol{u}^i_\lambda(.,y), \nabla^{\widetilde{\sigma}}\boldsymbol{\varphi}) = -(1/2)\big((\boldsymbol{a} + \boldsymbol{H})(.,y)e_i, D\boldsymbol{\varphi}\big)_2. \tag{21}$$

Having in mind to apply Proposition 6.1, we first prove

Lemma 6.2. *The mapping $y \mapsto \boldsymbol{b}_i(.,y) \in L^2(\Omega) \cap \mathbb{H}_{-1}$ is two times continuously differentiable in $\mathbb{H}_{-1}$, and the derivatives are bounded and Lipschitzian in $\mathbb{H}_1$.*

Proof: First note that for each $\boldsymbol{\varphi} \in \mathcal{C}$,

$$(\boldsymbol{b}_i(.,y), \boldsymbol{\varphi})_2 = -(1/2)\big((\boldsymbol{a} + \boldsymbol{H})(.,y)e_i, D\boldsymbol{\varphi}\big)_2.$$

From Assumption 2.4, we easily deduce that $\boldsymbol{b}_i(.,y) \in \mathbb{H}_{-1}$ and that the mapping $y \in \mathbb{R}^d \mapsto \boldsymbol{b}_i(.,y) \in \mathbb{H}_{-1}$ is bounded and Lipschitzian.

From Assumption 2.4 again, it is readily seen that the $\mathbb{H}_{-1}$ derivatives of $\boldsymbol{b}_i$ coincide, for $1 \leq k \leq d$, with the classical derivatives $\partial_{y_k}\boldsymbol{b}_i$ and

$$(\partial_{y_k}\boldsymbol{b}_i(.,y), \boldsymbol{\varphi})_2 = -(1/2)\big((\partial_{y_k}\boldsymbol{a} + \partial_{y_k}\boldsymbol{H})(.,y)e_i, D\boldsymbol{\varphi}\big)_2 \leq C\|\boldsymbol{\varphi}\|_1.$$

Since $\partial_{y_k}\boldsymbol{a}(\omega)$ and $\partial_{y_k}\boldsymbol{H}(\omega)$ are $(M, \tilde{\boldsymbol{a}}(\omega))$-controlled, the derivatives are bounded and Lipschitzian in $\mathbb{H}_1$. The same job can be carried out for the second order derivatives. Details are left to the reader. $\square$

From Proposition 6.1 (with $\boldsymbol{h} = 0$ and $\boldsymbol{f} = \boldsymbol{b}_i$), the mapping $y \mapsto \boldsymbol{u}_\lambda^i(.,y)$ is two times continuously differentiable in $\mathbb{H}_1$. We now investigate the asymptotic behavior of $\boldsymbol{u}_\lambda^i$ as well as its derivatives, as λ goes to zero.

Proposition 6.3. *For each fixed $y \in \mathbb{R}^d$ and $1 \leq i \leq d$, the family $(\nabla^{\widetilde{\sigma}} \boldsymbol{u}_\lambda^i(.,y))_\lambda$ converges to a limit $\widetilde{\boldsymbol{\xi}}_i(.,y) \in L^2(\Omega)^d$ as λ goes to 0. The same property holds for the derivatives, namely that the families $(\nabla^{\widetilde{\sigma}} \partial_{y_j} \boldsymbol{u}_\lambda^i)_\lambda$, $(\nabla^{\widetilde{\sigma}} \partial^2_{y_j y_k} \boldsymbol{u}_\lambda^i)_\lambda$ ($1 \leq i,j,k \leq d$) respectively converge to $\partial_{y_j} \widetilde{\boldsymbol{\xi}}_i(.,y)$, $\partial^2_{y_j y_j k} \widetilde{\boldsymbol{\xi}}_i(.,y)$ in $L^2(\Omega)^d$. Furthermore, we have*

$$\lambda |\boldsymbol{u}_\lambda^i(.,y)|_2^2 + \lambda |\partial_{y_j} \boldsymbol{u}_\lambda^i(.,y)|_2^2 + \lambda |\partial^2_{y_j y_k} \boldsymbol{u}_\lambda^i(.,y)|_2^2 \to 0, \quad \text{as } \lambda \text{ tends to } 0,$$

and, each function $\boldsymbol{g}_\lambda(.,y) = \boldsymbol{u}_\lambda^i(.,y), \partial_{y_k} \boldsymbol{u}_\lambda^i(.,y), \partial_{y_k y_l} \boldsymbol{u}_\lambda^i(.,y)$ satisfies the property:

$$\lambda |\boldsymbol{g}_\lambda(.,y)|_2^2 + \|\boldsymbol{g}_\lambda(.,y)\|_1^2 \leq C_{6.3} \tag{22}$$

$$\lambda |\boldsymbol{g}_\lambda(.,y+h) - \boldsymbol{g}_\lambda(.,y)|_2^2 + \|\boldsymbol{g}_\lambda(.,y+h) - \boldsymbol{g}_\lambda(.,y)\|_1^2 \leq C_{6.3} |h|^2 \tag{23}$$

for every $y, h \in \mathbb{R}^d$, where $C_{6.3}$ is a positive constant independent of $\lambda > 0$ and $y \in \mathbb{R}^d$.

Proof: The proof does not deeply differ from Proposition 4.3 in [13], but we nevertheless set it out because of its importance. From (19a) (note that $C_2 = 0$), we get $\lambda |\boldsymbol{u}_\lambda^i(.,y)|_2^2 + |\nabla^{\widetilde{\sigma}} \boldsymbol{u}_\lambda^i(.,y)|_2^2 \leq C$. Denote by $\widetilde{\boldsymbol{\xi}}_i(.,y) \in L^2(\Omega)^d$ a weak limit of the family $(\nabla^{\widetilde{\sigma}} \boldsymbol{u}_\lambda^i(.,y))_\lambda$ as λ goes to 0. Passing to the limit in (21), it is plain to see that $\forall \boldsymbol{\varphi} \in \mathcal{C}$

$$\boldsymbol{T}^y(\widetilde{\boldsymbol{\xi}}_i(.,y), \nabla^{\widetilde{\sigma}} \boldsymbol{\varphi}) = -(1/2)\big((\boldsymbol{a} + \boldsymbol{H})(.,y) e_i, D\boldsymbol{\varphi}\big)_2. \tag{24}$$

Since $\boldsymbol{T}^y$ is coercive on $\mathbb{D} \times \mathbb{D}$, this proves the uniqueness of the weak limit in $\mathbb{D}$. Gathering (21) and (24), we get

$$\lambda(\boldsymbol{u}_\lambda^i(.,y), \boldsymbol{\varphi})_2 + \boldsymbol{T}^y(\nabla^{\widetilde{\sigma}} \boldsymbol{u}_\lambda^i(.,y), \nabla^{\widetilde{\sigma}} \boldsymbol{\varphi}) = \boldsymbol{T}^y(\widetilde{\boldsymbol{\xi}}_i(.,y), \nabla^{\widetilde{\sigma}} \boldsymbol{\varphi}). \tag{25}$$

Choosing $\boldsymbol{u}_\lambda^i(.,y) = \boldsymbol{\varphi}$ yields:

$$\lambda |\boldsymbol{u}_\lambda^i(.,y)|_2^2 + \boldsymbol{T}^y\big(\nabla^{\widetilde{\sigma}} \boldsymbol{u}_\lambda^i(.,y), \nabla^{\widetilde{\sigma}} \boldsymbol{u}_\lambda^i(.,y)\big) \leq \boldsymbol{T}^y\big(\widetilde{\boldsymbol{\xi}}_i(.,y), \widetilde{\boldsymbol{\xi}}_i(.,y)\big) + \epsilon(\lambda),$$

where the function $\epsilon(\lambda)$ exactly matches $\boldsymbol{T}^y\big(\widetilde{\boldsymbol{\xi}}_i(.,y), \nabla^{\widetilde{\sigma}} \boldsymbol{u}_\lambda^i(.,y) - \widetilde{\boldsymbol{\xi}}_i(.,y)\big)$ and thus converges to 0 as λ goes to 0. Hence $\limsup_{\lambda \to 0} \boldsymbol{T}^y\big(\nabla^{\widetilde{\sigma}} \boldsymbol{u}_\lambda^i(.,y), \nabla^{\widetilde{\sigma}} \boldsymbol{u}_\lambda^i(.,y)\big) \leq \boldsymbol{T}^y\big(\widetilde{\boldsymbol{\xi}}_i(.,y), \widetilde{\boldsymbol{\xi}}_i(.,y)\big)$. Denote by $\boldsymbol{T}^S$ the symmetric part of $\boldsymbol{T}^y$

$$\boldsymbol{T}^S(\boldsymbol{\varphi}, \boldsymbol{\psi}) = (1/2)\big[\boldsymbol{T}^y(\boldsymbol{\varphi}, \boldsymbol{\psi}) + \boldsymbol{T}^y(\boldsymbol{\psi}, \boldsymbol{\varphi})\big], \quad \boldsymbol{\varphi}, \boldsymbol{\psi} \in \mathbb{D}.$$

From Assumption 2.4 and the antisymmetry of $\boldsymbol{H}$, we have

$$M^{-1}(\widetilde{\boldsymbol{\sigma}}^* D\boldsymbol{\varphi}, \widetilde{\boldsymbol{\sigma}}^* D\boldsymbol{\varphi})_2 \leq \boldsymbol{T}^S(\nabla^{\widetilde{\boldsymbol{\sigma}}} \boldsymbol{\varphi}, \nabla^{\widetilde{\boldsymbol{\sigma}}} \boldsymbol{\varphi}) \leq M(\widetilde{\boldsymbol{\sigma}}^* D\boldsymbol{\varphi}, \widetilde{\boldsymbol{\sigma}}^* D\boldsymbol{\varphi})_2, \quad \boldsymbol{\varphi} \in \mathcal{C}.$$

By density arguments, the quadratic form associated to $\boldsymbol{T}^S$ defines a norm on $\mathbb{D}$ equivalent to the canonical inner product. Moreover, we have just proved that the family $(\nabla^{\widetilde{\sigma}}\boldsymbol{u}^i_\lambda(.,y))_\lambda$ is weakly convergent in $\mathbb{D}$ to $\widetilde{\boldsymbol{\xi}}_i(.,y)$ and $\limsup_{\lambda\to 0}\boldsymbol{T}^S\big(\nabla^{\widetilde{\sigma}}\boldsymbol{u}^i_\lambda(.,y),\nabla^{\widetilde{\sigma}}\boldsymbol{u}^i_\lambda(.,y)\big)\leq \boldsymbol{T}^S\big(\widetilde{\boldsymbol{\xi}}_i(.,y),\widetilde{\boldsymbol{\xi}}_i(.,y)\big)$. Thus the convergence is strong with respect to the norm on $\mathbb{D}$ associated to $\boldsymbol{T}^S$, and consequently $(\nabla^{\widetilde{\sigma}}\boldsymbol{u}^i_\lambda(.,y))_\lambda$ strongly converges in $(L^2(\Omega))^d$ to $\widetilde{\boldsymbol{\xi}}_i(.,y)$. From this together with (25), we get

$$\lambda|\boldsymbol{u}^i_\lambda(.,y)|_2^2+|\nabla^{\widetilde{\sigma}}\boldsymbol{u}^i_\lambda(.,y)-\widetilde{\boldsymbol{\xi}}_i(.,y)|_2^2\to 0 \text{ as } \lambda\to 0.$$

This proves the first part of the statement for the function $\boldsymbol{u}^i_\lambda(.,y)$. The second part results from Proposition 6.1, statements (19a) and (19b) (with $C_2=0$). The same job can be carried out for the successive derivatives of $\boldsymbol{u}^i_\lambda(.,y)$ up to order 2. □

7 Dynamics of the process X^ε. Preliminary results

Notations. *All the results of this section are valid for any value of the parameter ε. However, to simplify the notations, we choose $\varepsilon=1$ and thus remove the parameter ε from the notations. So the process X stands for the process X^ε defined by* (7). *Finally we denote by $\mathbb{P}_V$ the probability measure $e^{-2V(y)}\,dy\otimes d\mu$ on $\Omega\times\mathbb{R}^d$ and by $\mathbb{M}_V$ the coresponding expectation.*

This section is devoted to the study of the $\Omega\times\mathbb{R}^d$-valued process $(\tau_X\omega,X)$, such as its invariant distribution and the Itô formula. Since these properties are more easily established when the process X possesses regularizing properties, namely that the diffusion coefficient $\boldsymbol{a}$ is uniformly elliptic, most of the following proofs are carried out through vanishing viscosity methods, that is, in considering a family of non-degenerate diffusion processes that converges to X.

Invariant distribution. Let us introduce a standard d-dimensional Brownian motion $\tilde{B}$ independent of B. For each fixed $(\omega,n)\in\Omega\times\bar{\mathbb{N}}^*$ and for any $x\in\mathbb{R}^d$, we define the Itô process X^n as the solution of the SDE (with the convention $n^{-1}=0$ if $n=\infty$)

$$X^n_t=x+\int_0^t(b+c-n^{-1}\partial_yV)(\omega,X^n_r,X^n_r)\,dr+\int_0^t\sigma(\omega,X^n_r,X^n_r)\,dB_r+(n/2)^{-1/2}\tilde{B}_t.$$

Note that, for $n=\infty$, X^∞ coincides with the process X. For $n\in\bar{\mathbb{N}}^*$, the process X^n defines a continuous semigroup P^n on $C_b(\mathbb{R}^d)$ (continuous bounded functions). Its generator $\mathcal{L}^n$ coincides on $C^2(\mathbb{R}^d)$ with

$$\mathcal{L}^n=\frac{1}{2}e^{2V(x)}\sum_{i,j}\partial_{x_i}\big(e^{-2V(x)}(a+H+n^{-1}\mathrm{Id})_{ij}(\omega,x,x)\partial_{x_j}\cdot\big). \tag{26}$$

For $n\in\mathbb{N}^*$, it is well-known that the distribution of X^n_t ($t>0$) admits a density $p^n(\omega,t,x,\cdot)$ with respect to the Lebesgue measure (cf. [14, Sect. II.2]), which is bounded from above

by a constant C that only depends on Λ, n, t. Thus the semigroup associated to X^n ($n \in \mathbb{N}^*$) continuously extends to $L^2(\mathbb{R}^d, e^{-2V(x)}\,dx)$. Let us denote by $(\mathcal{L}^n)^*$ the adjoint of $\mathcal{L}^n$ in $L^2(\mathbb{R}^d, e^{-2V(x)}\,dx)$, which coincides on $C^2(\mathbb{R}^d)$ with

$$(\mathcal{L}^n)^* = \frac{1}{2}e^{2V(x)}\sum_{i,j}\partial_{x_i}\big(e^{-2V(x)}(a - H + n^{-1}\mathrm{Id})_{ij}(\omega, x, x)\partial_{x_j}\,\cdot\,\big). \tag{27}$$

Now, for $\varphi, \psi \in C_c^\infty(\mathbb{R}^d)$, let us compute $\int_{\mathbb{R}^d}\mathcal{L}^n P_t^n\varphi(x)\psi(x)e^{-2V(x)}\,dx$. From [8], $P_t^n\varphi \in C^2(\mathbb{R}^d)$ so that $\mathcal{L}^n P_t^n\varphi$ can be computed with the help of (26). By integrating by parts, we obtain

$$\int_{\mathbb{R}^d}\mathcal{L}^n P_t^n\varphi(x)\psi(x)e^{-2V(x)}\,dx = \int_{\mathbb{R}^d}P_t^n\varphi(x)(\mathcal{L}^n)^*\psi(x)e^{-2V(x)}\,dx. \tag{28}$$

Moreover, we have $\mathcal{L}^n P_t^n\varphi = P_t^n\mathcal{L}^n\varphi \in C_b(\mathbb{R}^d)$. Choose now a function $\varrho \in C_c^\infty(\mathbb{R}^d)$ that matches 1 over the ball $B(0;1)$. Define $\psi_m(x) = \varrho(x/m)$. It is readily seen that the sequence $(\mathcal{L}^n\psi_m)_m$ is bounded in $L^\infty(\mathbb{R}^d)$ and uniformly converges to 0 on the compact subsets of $\mathbb{R}^d$. Thus, choosing $\psi = \psi_m$ in (28), and passing to the limit as m goes to ∞, we get

$$\forall\varphi \in C_c^\infty(\mathbb{R}^d), \quad \int_{\mathbb{R}^d}\mathcal{L}^n P_t^n\varphi(x)e^{-2V(x)}\,dx = 0. \tag{29}$$

In particular, for any $\varphi \in C_c^\infty(\mathbb{R}^d)$, $\int_{\mathbb{R}^d}P_t^n\varphi(x)e^{-2V(x)}\,dx = \int_{\mathbb{R}^d}\varphi(x)e^{-2V(x)}\,dx$, in such a way that, by density arguments, the probability measure $e^{-2V(x)}\,dx$ is invariant for the process X^n ($n \geq 1$). Then classical arguments of SDE theory ensure that the sequence of processes $(X^n)_n$ converges in law in $C([0,T];\mathbb{R}^d)$ to the process X as n goes to ∞. We deduce that $\int_{\mathbb{R}^d}P_t\varphi(x)e^{-2V(x)}\,dx = \int_{\mathbb{R}^d}\varphi(x)e^{-2V(x)}\,dx$ holds for $\varphi \in C_b(\mathbb{R}^d)$. The semigroup associated to X thus extends to $L^p(\mathbb{R}^d; e^{-2V(x)}\,dx)$ for $p \geq 1$ and the probability measure $e^{-2V(x)}\,dx$ is also invariant for this semigroup.

Finally, for each $\boldsymbol{\varphi} \in C_b(\Omega \times \mathbb{R}^d)$ (i.e. for each fixed $\omega \in \Omega$, the function $x \mapsto \boldsymbol{\varphi}(\tau_x\omega, x)$ is continuous and bounded by a constant independent of ω) and $n \geq 0$, we deduce from the previous remarks and the invariance of the measure μ under space translations that

$$\bar{\mathbb{E}}[\boldsymbol{\varphi}(\tau_{X_t^n}\omega, X_t^n)] = \mathbb{M}_V[\boldsymbol{\varphi}(\tau_x\omega, x)] = \mathbb{M}_V[\boldsymbol{\varphi}(\omega, x)], \tag{30}$$

so that the mapping $\boldsymbol{\varphi} \in C_b(\Omega \times \mathbb{R}^d) \mapsto P_t^n(\boldsymbol{\varphi}) = \mathbb{E}_x[\boldsymbol{\varphi}(\tau_{X_t^n}\omega, X_t^n)]$ continuously extends to $L^p(\Omega \times \mathbb{R}^d; \mathbb{P}_V)$ for any $p \geq 1$ and (30) holds for $\boldsymbol{\varphi} \in L^p(\Omega \times \mathbb{R}^d; \mathbb{P}_V)$.

Itô's formula. We now aim at establishing the Itô formula to the process $(\tau_X\omega, X)$ and to the function $(x, y) \mapsto u_\lambda(\omega, x, y)$, where $\boldsymbol{u}_\lambda$ is the solution of the resolvent equation (18), with functions $\boldsymbol{h}(., y)$ and $\boldsymbol{f}(., y)$ satisfying the assumptions of Proposition 6.1. This latter proposition describes the regularity of u_λ with respect to the variable y. Due to the possible

degeneracies of $\boldsymbol{\sigma}$, the difficulty actually lies in the regularity with respect to the parameter $x \in \mathbb{R}^d$. To apply the Itô formula and get round technical difficulties, we use viscosity methods again, namely that we look at the operator $\lambda - \boldsymbol{L}^y - n^{-1}\boldsymbol{\Delta}$ for $n \in \mathbb{N}^*$. Obviously, there is no difficulty in solving the corresponding resolvent equation with the techniques used in Section 6 (it suffices to replace $\boldsymbol{a}$ by $\boldsymbol{a} + n^{-1}\mathrm{Id}$ and to choose $\tilde{\boldsymbol{a}} = \mathrm{Id}$)

$$\lambda \boldsymbol{u}_\lambda^{(n)}(\cdot, y) - \big(\boldsymbol{L}^y + n^{-1}\boldsymbol{\Delta}\big)\boldsymbol{u}_\lambda^{(n)}(\cdot, y) = \boldsymbol{h}(\cdot, y) + \boldsymbol{f}(., y). \tag{31}$$

The strategy then consists in applying the Itô formula in the non-zero viscosity setting and then in letting n tend to ∞. Thanks to the regularizing parameter $n \in \mathbb{N}^*$, the Itô formula holds in the non-zero viscosity setting (cf [13, Sect. 5]). The following formula thus holds

$$\begin{aligned} du_\lambda^{(n)}(X_t^n, X_t^n) =& (\lambda u_\lambda^{(n)} - h - f)(X_t^n, X_t^n)\,dt + [c - n^{-1}\partial_y \overline{V}] \cdot Du_\lambda^{(n)}(X_t^n, X_t^n)\,dt \\ &+ (\nabla^{\sigma(.,y)} u_\lambda^{(n)})^*(X_t^n, X_t^n)\,dB_t + n^{-1/2}(Du_\lambda^{(n)})(X_t^n, X_t^n)\,d\widetilde{B}_t \\ &+ b\partial_y u_\lambda^{(n)}(X_t^n, X_t^n)\,dt + [c - n^{-1}\partial_y V] \cdot \partial_y u_\lambda^{(n)}(X_t^n, X_t^n)\,dt \\ &+ (\partial_y u_\lambda^{(n)})^* \sigma(X_t^n, X_t^n)\,dB_t + n^{-1/2}(\partial_y u_\lambda^{(n)})(X_t^n, X_t^n)\,d\widetilde{B}_t \\ &+ (1/2)\mathrm{trace}([a + n^{-1}Id]\partial^2_{yy} u_\lambda^{(n)})(X_t^n, X_t^n)\,dt \\ &+ \mathrm{trace}([a + n^{-1}Id]D\partial_y u_\lambda^{(n)})(X_t^n, X_t^n)\,dt. \end{aligned} \tag{32}$$

Having in mind to let n tend to ∞ in (32), let us now describe the behavior of $\boldsymbol{u}_\lambda^n$ as n tends to ∞. We first claim:

Proposition 7.1.

$$\lim_{n\to\infty} \Big[|\boldsymbol{u}_\lambda^{(n)}(., y) - \boldsymbol{u}_\lambda(., y)|_2 + \|\boldsymbol{u}_\lambda^{(n)}(., y) - \boldsymbol{u}_\lambda(., y)\|_1 + n^{-1}|D\boldsymbol{u}_\lambda^{(n)}(., y)|_2^2 \Big] = 0, \tag{33}$$

and that there exists a constant D_{34} (independent of n and $y \in \mathbb{R}^d$) such that

$$\begin{aligned} |\boldsymbol{u}_\lambda^{(n)}(\cdot, y+h) - \boldsymbol{u}_\lambda^{(n)}(\cdot, y)|_2^2 + \|\boldsymbol{u}_\lambda^{(n)}(\cdot, y+h) - \boldsymbol{u}_\lambda^{(n)}(\cdot, y)\|_1^2 \\ + n^{-1}|D\boldsymbol{u}_\lambda^{(n)}(\cdot, y+h) - D\boldsymbol{u}_\lambda^{(n)}(\cdot, y)|_2^2 \le D_{34}|h|^2. \end{aligned} \tag{34}$$

Moreover, the same properties hold for the sequences $(\partial_{y_k}\boldsymbol{u}_\lambda^{(n)})_n, (\partial^2_{y_k y_l}\boldsymbol{u}_\lambda^{(n)})_n$ and their corresponding limits $(\partial_{y_k}\boldsymbol{u}_\lambda)_n, (\partial^2_{y_k y_l}\boldsymbol{u}_\lambda)_n$, for $1 \le k, l \le d$.

Proof. Since the proofs of (33) and (34) can be adapted from the proof of Proposition 6.3, we just set out the guiding line of (33).

To clarify the notations, we forget for a while the dependence on the parameter y. First multiply (31) by $\boldsymbol{u}_\lambda^{(n)}$ and integrate with respect to the measure μ so as to obtain the estimate:

$$\lambda|\boldsymbol{u}_\lambda^{(n)}|_2^2 + |\nabla^{\widetilde{\sigma}}\boldsymbol{u}_\lambda^{(n)}|_2^2 + n^{-1}|D\boldsymbol{u}_\lambda^{(n)}|_2^2 \le C$$

for some constant C only depending on $|\boldsymbol{h}|_2^2/\lambda$ and $\|\boldsymbol{f}\|_{-1}^2$. From this estimate, we deduce that the family $(n^{-1}D\boldsymbol{u}_\lambda^{(n)})_n$ strongly converges to 0 in $(L^2(\Omega))^d$ as $n \to \infty$ and that, up to extracting a subsequence, the family $(\boldsymbol{u}_\lambda^{(n)})_n$ weakly converges in $\mathbb{H}_1$ as $n \to \infty$. Multiply once again (31) by a test function $\boldsymbol{\varphi} \in \mathcal{C}$, integrate with respect to the measure μ and then pass to the limit as $n \to \infty$ to identity the weak limit in $\mathbb{H}_1$ as being necessarily equal to $\boldsymbol{u}_\lambda$. So the whole family $(\boldsymbol{u}_\lambda^{(n)})_n$ is weakly convergent in $\mathbb{H}_1$ (not up to a subsequence). It just remains to prove that the convergence actually holds in the strong sense. We can integrate (31) and (18) against a test function $\boldsymbol{\varphi} \in \mathcal{C}$. Since the right-hand sides of (31) and (18) coincide, this yields:

$$\lambda(\boldsymbol{u}_\lambda^{(n)}, \boldsymbol{\varphi})_2 + \boldsymbol{T}^y(\nabla^{\widetilde{\sigma}}\boldsymbol{u}_\lambda^{(n)}, \nabla^{\widetilde{\sigma}}\boldsymbol{\varphi}) + n^{-1}(D\boldsymbol{u}_\lambda^{(n)}, D\boldsymbol{\varphi})_2 = \lambda(\boldsymbol{u}_\lambda, \boldsymbol{\varphi})_2 + \boldsymbol{T}^y(\nabla^{\widetilde{\sigma}}\boldsymbol{u}_\lambda, \nabla^{\widetilde{\sigma}}\boldsymbol{\varphi}).$$

Choose $\boldsymbol{\varphi} = \boldsymbol{u}_\lambda^{(n)}$ and pass to the limit as $n \to \infty$ and get

$$\lim_{n\to\infty}\Big(\lambda|\boldsymbol{u}_\lambda^{(n)}|_2^2 + \boldsymbol{T}^y(\nabla^{\widetilde{\sigma}}\boldsymbol{u}_\lambda^{(n)}, \nabla^{\widetilde{\sigma}}\boldsymbol{u}_\lambda^{(n)}) + n^{-1}|D\boldsymbol{u}_\lambda^{(n)}|_2^2\Big) = \lambda|\boldsymbol{u}_\lambda|_2^2 + \boldsymbol{T}^y(\nabla^{\widetilde{\sigma}}\boldsymbol{u}_\lambda, \nabla^{\widetilde{\sigma}}\boldsymbol{u}_\lambda).$$

As in Proposition 6.3, this is sufficient to establish the strong convergence of $(\boldsymbol{u}_\lambda^{(n)})_n$ in $\mathbb{H}_1$ and, consequently, the convergence $n^{-1}|D\boldsymbol{u}_\lambda^{(n)}|_2^2 \to 0$ as $n \to \infty$. $\square$

We are now in position to conclude. Going through formula (32), we are faced with functionals of type $\int_t^s g_n(X_r^n, X_r^n)\, dr$ (concerning the martingale terms, it suffices to work on their quadratic variations), where $\mathbb{M}_V[|\boldsymbol{g}_n - \boldsymbol{g}_0|] \to 0$ as n tends to ∞ and

$$\forall (y,h) \in \mathbb{R}^d \times \mathbb{R}^d, \quad |\boldsymbol{g}_n(.,y+h) - \boldsymbol{g}_n(.,y)|_2 \leq C|h| \tag{35}$$

where the constant C depends neither on $n \in \mathbb{N}$ nor $y, h \in \mathbb{R}^d$. From Lemma 7.3 below, we prove the convergence of the functional towards $\int_t^s g_0(X_r, X_r)\, dr$ in $\bar{\mathbb{P}}$-probability and as a consequence the

Theorem 7.2. *Let $\boldsymbol{h}, \boldsymbol{f}$ be two functions satisfying the assumptions of Proposition 6.1. Let $\boldsymbol{u}_\lambda$ be the solution of the resolvent equation:*

$$\lambda\boldsymbol{u}_\lambda(\cdot,y) - \boldsymbol{L}^y\boldsymbol{u}_\lambda(\cdot,y) = \boldsymbol{h}(\cdot,y) + \boldsymbol{f}(\cdot,y).$$

Then the following Itô formula holds (we reintroduce the parameter ε):

$$\begin{aligned}\varepsilon du_\lambda(\overline{X}_t^\varepsilon, X_t^\varepsilon) =& \varepsilon^{-1}(\lambda u_\lambda - h - f)(\overline{X}_t^\varepsilon, X_t^\varepsilon)\, dt + c\cdot Du_\lambda(\overline{X}_t^\varepsilon, X_t^\varepsilon)\, dt \\ &+ (\nabla^{\sigma(.,y)}u_\lambda)^*(\overline{X}_t^\varepsilon, X_t^\varepsilon)\, dB_t + b\partial_y u_\lambda(\overline{X}_t^\varepsilon, X_t^\varepsilon)\, dt \\ &+ \varepsilon(\partial_y u_\lambda)^*\sigma(\overline{X}_t^\varepsilon, X_t^\varepsilon)\, dB_t + \varepsilon c\cdot\partial_y u_\lambda(\overline{X}_t^\varepsilon, X_t^\varepsilon)\, dt \\ &+ (\varepsilon/2)\mathrm{trace}(a\partial_{yy}^2 u_\lambda)(\overline{X}_t^\varepsilon, X_t^\varepsilon)\, dt + \mathrm{trace}(aD\partial_y u_\lambda)(\overline{X}_t^\varepsilon, X_t^\varepsilon)\, dt.\end{aligned}$$

Lemma 7.3. *Consider a sequence of functions $\boldsymbol{g}_n \in L^1(\Omega \times \mathbb{R}^d; \mathbb{P}_V)$ $(n \geq 0)$ such that $\mathbb{M}_V[|\boldsymbol{g}_n - \boldsymbol{g}_0|] \to 0$ as $n \to \infty$ and for any $(y,h) \in \mathbb{R}^d \times \mathbb{R}^d$, $|\boldsymbol{g}_n(.,y+h) - \boldsymbol{g}_n(.,y)|_2 \leq C|h|$ for some constant C that depends neither on n nor $y, h \in \mathbb{R}^d$.*

Then $\bar{\mathbb{E}}[|g_n(X_r^n, X_r^n) - g_0(X_r, X_r)|] \to 0$ as $n \to 0$.

Proof: First, suppose that $\boldsymbol{g}_0$ is bounded. Let us consider a smooth mollifier $p : \mathbb{R}^d \to \mathbb{R}$ and $\varrho \in C_c^\infty(\mathbb{R}^d)$ such that $\varrho = 1$ over the ball $B(0;1)$. We define for $m, q \geq 1$, $p_m(\cdot) = m^d p(m\,\cdot)$, $\varrho_q(\cdot) = \varrho(\cdot/q)$ and $\boldsymbol{g}_0^{m,q}(\omega, x) = \int_{\mathbb{R}^d} \boldsymbol{g}_0(\tau_{-x'}\omega, x')\varrho_q(x')p_m(x-x')\,dx'$. Then, from (30),

$$\begin{aligned}\bar{\mathbb{E}}[|g_n(X_r^n, X_r^n) - g_0(X_r, X_r)|] \leq & \mathbb{M}_V[|\boldsymbol{g}_n - \boldsymbol{g}_0|] + 2\mathbb{M}_V[|\boldsymbol{g}_0^{m,q} - \boldsymbol{g}_0|] \\ & + \bar{\mathbb{E}}[|g_0^{m,q}(X_r^n, X_r^n) - g_0^{m,q}(X_r, X_r)|].\end{aligned}$$

With classical convolution techniques, we can prove that m, q can be chosen large enough to make the term $2\mathbb{M}_V[|\boldsymbol{g}_0^{m,q} - \boldsymbol{g}_0|]$ small. Then, from the Lipschitz regularity of the coefficients (Assumption 2.2), the classical theory of SDEs ensures that $\mathbb{E}_x[\sup_{0\leq t\leq T}|X_t^n - X_t|^2] \leq n^{-1}D$ for some constant D that only depends on M, Λ and T. For each fixed $m, q \geq 1$ and $\omega \in \Omega$, the function $x \mapsto g_0^{m,q}(x,x)$ is continuous with compact support so that $\int_{\mathbb{R}^d} \mathbb{E}_x[|g_0^{m,q}(X_r^n, X_r^n) - g_0^{m,q}(X_r, X_r)|]e^{-2V(x)}\,dx \to 0$ as $n \to \infty$. Then, the Lebesgue theorem ($g_0^{m,q}$ is bounded independently from ω) proves that $\bar{\mathbb{E}}[|g_0^{m,q}(X_r^n, X_r^n) - g_0^{m,q}(X_r, X_r)|]$ converges to 0 as n goes to ∞. Therefore, n can be chosen large enough to make this latter term small. Finally, from the assumptions of the lemma, even if it means considering larger n, the term $\mathbb{M}_V[|\boldsymbol{g}_n - \boldsymbol{g}_0|]$ is small too. The proof is then easily completed in the case when $\boldsymbol{g}_0$ is bounded.

If $\boldsymbol{g}_0$ is not bounded, it suffices to consider for $n \geq 0$ and $R > 0$, $\boldsymbol{g}_n^R = \max(-R; \min(\boldsymbol{g}_n; R))$. It is readily checked that the sequence $(\boldsymbol{g}_n^R)_n$ still satisfies all the assumptions of the lemma in such a way that $\bar{\mathbb{E}}[|g_n^R(X_r^n, X_r^n) - g_0^R(X_r, X_r)|] \to 0$ as $n \to 0$, for each fixed $R > 0$. Then, from (30), $\bar{\mathbb{E}}[|g_n^R(X_r^n, X_r^n) - g_n(X_r^n, X_r^n)|] \leq \mathbb{M}_V[|\boldsymbol{g}_n^R - \boldsymbol{g}_n|]$ and

$$\lim_{R\to\infty}\lim_{n\to\infty} \mathbb{M}_V[|\boldsymbol{g}_n^R - \boldsymbol{g}_n|] = \lim_{R\to\infty} \mathbb{M}_V[|\boldsymbol{g}_0^R - \boldsymbol{g}_0|] = 0.$$

Since we have

$$\begin{aligned}\bar{\mathbb{E}}[|g_n(X_r^n, X_r^n) - g_0(X_r^n, X_r^n)|] \leq & \bar{\mathbb{E}}[|g_n^R(X_r^n, X_r^n) - g_n(X_r^n, X_r^n)|] \\ & + \bar{\mathbb{E}}[|g_n^R(X_r^n, X_r^n) - g_0^R(X_r, X_r)|] + \bar{\mathbb{E}}[|g_0^R(X_r, X_r) - g_0(X_r^n, X_r^n)|],\end{aligned}$$

the proof is then easily completed in this case too. □

8 Asymptotic Theorems

Classical ergodic theorem. In this section, we aim at exploiting the asymptotic properties of the process X^ε, more precisely Assumption 2.5, in order to describe the asymptotic behavior

of functionals of type $\int_0^t \Psi(\overline{X}^\varepsilon_r, X^\varepsilon_r)\,dr$ for a suitable locally stationary random field Ψ. The classical ergodic theory leads us to guess that the local ergodicity assumption 2.5 makes this functional average with respect to its first variable. More precisely,

Theorem 8.1. (Ergodic Theorem) *Let us consider* $\boldsymbol{\Psi} : \Omega\times\mathbb{R}^d \to \mathbb{R}$ *such that* $\mathbb{M}_V[|\boldsymbol{\Psi}|] < +\infty$. *Denoting* $\overline{\Psi}(y) = \mathbb{M}[\boldsymbol{\Psi}(\cdot, y)]$, *the following convergence holds:*

$$\bar{\mathbb{E}}^\varepsilon\big[\sup_{0\leq s\leq t}|\int_0^s \Psi(\overline{X}^\varepsilon_r, X^\varepsilon_r)\,dr - \int_0^s \overline{\Psi}(X^\varepsilon_r)\,dr|^2\big] \xrightarrow[\varepsilon\to 0]{} 0. \tag{36}$$

Proof: This result can be proved in the same way as [13, Th. 6.1]. The only difference consists in establishing: $\boldsymbol{g} \in \mathrm{Dom}(\boldsymbol{L}^y) \subset \mathbb{H}_1$ and $\boldsymbol{L}^y\boldsymbol{g} = 0$ implies that $\boldsymbol{g}$ is constant μ almost surely. In the uniformly elliptic setting, it turns out that the derivatives $D_i\boldsymbol{g}$ reduce to 0 and, as a consequence, $\boldsymbol{g}$ is constant. In the degenerate framework, we need to use Assumption 2.5 as follows. From Assumption 2.4, $\|\boldsymbol{g}\|_1^2 \leq M\|\boldsymbol{g}\|_{1,y}^2 = -(\boldsymbol{g}, \boldsymbol{L}^y\boldsymbol{g})_2 = 0$. In particular, $\mathcal{B}^{S^y}(\boldsymbol{g}, \cdot) = 0$. Hence $\boldsymbol{g} \in \mathrm{Dom}(\boldsymbol{S}^y)$ and $\boldsymbol{S}^y\boldsymbol{g} = 0$. Thus $\boldsymbol{g}$ is constant (Assumption 2.5). $\square$

Asymptotic theorem for highly oscillating functionals. Theorem 8.1 describes the asymptotic behavior of functionals of type $\int_0^t \Psi(\overline{X}^\varepsilon_r, X^\varepsilon_r)\,dr$ in order to pass to the limit in (7). However, as explained in [13], additional difficulties arise in the random setting in comparison with the periodic one. In particular, we must describe the asymptotic behavior of the functional $\int_0^t \Psi_\varepsilon(\overline{X}^\varepsilon_r, X^\varepsilon_r)\,dr$ for a family $(\boldsymbol{\Psi}_\varepsilon)_\varepsilon$ that need not be convergent in $L^1(\Omega\times\mathbb{R}^d;\mathbb{P}_V)$ but satisfies a sort of uniform Poincaré inequality. Unlike [13, Theorem 6.3], technical difficulties due to the degeneracy of the diffusion coefficient $\boldsymbol{a}$ occur. In particular, because of the lack of Aronson type estimates, the tightness of the process X^ε is not obvious. To prove this tightness, all asymptotic convergences need be established in $C([0,T];\mathbb{R}^d)$ (note the $\sup$ in (38)). This is one of the main difficulty of Theorem 8.2 below in comparison with the uniformly elliptic setting (see [13, Theorem 6.3]). The strategy consists in expressing $\int_0^t \Psi_\varepsilon(\overline{X}^\varepsilon_r, X^\varepsilon_r)\,dr$ as the sum of two martingales thanks to time reversal arguments, and then in using the Doob inequality. The Poincaré inequality (37) ensures that the martingales possess suitable asymptotic properties.

Theorem 8.2. (Ergodic theorem II) *Let us consider, for each* $\varepsilon > 0$, *a function* $\boldsymbol{\Psi}_\varepsilon \in L^2(\Omega\times\mathbb{R}^d;\mathbb{P}_V)$ *satisfying the following Poincaré inequality: for any* $\boldsymbol{\varphi}(\omega, x) = \boldsymbol{\chi}(\omega)\varrho(x)$, $(\boldsymbol{\chi}, \varrho) \in \mathcal{C}\times C_c^\infty(\mathbb{R}^d)$,

$$\mathbb{M}_V[\boldsymbol{\Psi}_\varepsilon\boldsymbol{\varphi}] \leq C_\varepsilon\big(\mathbb{M}_V[|\boldsymbol{\sigma}^*(D+\varepsilon\partial_y)\boldsymbol{\varphi}|^2]\big)^{1/2}, \tag{37}$$

for some family $(C_\varepsilon)_{\varepsilon>0}$ *satisfying* $\varepsilon C_\varepsilon \to 0$ *as* $\varepsilon \to 0$. *Then*

$$\bar{\mathbb{E}}^\varepsilon\big[\sup_{0\leq s\leq t}|\int_0^s \Psi_\varepsilon(\overline{X}^\varepsilon_r, X^\varepsilon_r)\,dr|^2\big] \xrightarrow[\varepsilon\to 0]{} 0. \tag{38}$$

Proof: In what follows, we say that $\boldsymbol{\varphi} \in \mathcal{C}_\Pi$ if $\boldsymbol{\varphi}(\omega, y) = \boldsymbol{\chi}(\omega)\varrho(y)$, where $(\boldsymbol{\chi}, \varrho) \in \mathcal{C} \times C_c^\infty(\mathbb{R}^d)$. We aim at constructing, as prescribed in Section 5, the unbounded operators on $L^2(\Omega \times \mathbb{R}^d; \mathbb{P}_V)$ that coincide on $\mathcal{C}_\Pi$ for $n \in \bar{\mathbb{N}}^*$ with (here we use the convention $n^{-1} = 0$ if $n = \infty$)

$$S^{n,\varepsilon}\boldsymbol{\varphi} =(1/2)e^{2V} \sum_{i,j=1,\dots,d}(D_i+\varepsilon\partial_{y_i})\big[e^{-2V}(\boldsymbol{a}+n^{-1}\mathrm{Id})_{ij}(D_j+\varepsilon\partial_{y_j})\boldsymbol{\varphi}\big], \tag{39}$$

$$L^{n,\varepsilon}\boldsymbol{\varphi} =(1/2)e^{2V} \sum_{i,j=1,\dots,d}(D_i+\varepsilon\partial_{y_i})\big[e^{-2V}(\boldsymbol{a}+\boldsymbol{H}+n^{-1}\mathrm{Id})_{ij}(D_j+\varepsilon\partial_{y_j})\boldsymbol{\varphi}\big]. \tag{40}$$

For $\varepsilon > 0$, $n \in \bar{\mathbb{N}}^*$ and $\boldsymbol{\varphi}, \boldsymbol{\psi} \in \mathcal{C}_\Pi$, we define the corresponding bilinear forms

$$\langle \boldsymbol{\varphi}, \boldsymbol{\psi}\rangle_{n,\varepsilon} = (1/2)\mathbb{M}_V\big[(D\boldsymbol{\varphi}+\varepsilon\partial_y\boldsymbol{\varphi})^*(\boldsymbol{a}+n^{-1}\mathrm{Id})(D\boldsymbol{\psi}+\varepsilon\partial_y\boldsymbol{\psi})\big], \tag{41}$$

$$B_{n,\varepsilon}(\boldsymbol{\varphi}, \boldsymbol{\psi}) = (1/2)\mathbb{M}_V\big[(D\boldsymbol{\varphi}+\varepsilon\partial_y\boldsymbol{\varphi})^*(\boldsymbol{a}+\boldsymbol{H}+n^{-1}\mathrm{Id})(D\boldsymbol{\psi}+\varepsilon\partial_y\boldsymbol{\psi})\big]. \tag{42}$$

Clearly, $\langle\cdot,\cdot\rangle_{n,\varepsilon}$ is positive symmetric (denote by $\|\cdot\|_{n,\varepsilon}$ the corresponding seminorm). Note that, for each fixed $\varepsilon > 0$, the seminorms $(\|\cdot\|_{n,\varepsilon})_{n\in\mathbb{N}^*}$ are all equivalent. Moreover, for $n \in \bar{\mathbb{N}}^*$, $\|\boldsymbol{\varphi}\|_{n,\varepsilon}^2 \le B_{n,\varepsilon}(\boldsymbol{\varphi},\boldsymbol{\varphi})$ and $B_{n,\varepsilon}(\boldsymbol{\varphi},\boldsymbol{\psi}) \le 2M^2\|\boldsymbol{\varphi}\|_{n,\varepsilon}\|\boldsymbol{\psi}\|_{n,\varepsilon}$ for any $\boldsymbol{\varphi}, \boldsymbol{\psi} \in \mathcal{C}_\Pi$ (see Assumption 2.4). From Section 5, we can define

$$\begin{aligned}(\mathbb{H}_{n,\varepsilon}, S^{n,\varepsilon}, \mathrm{Dom}(S^{n,\varepsilon}), (G_\lambda^{S,n,\varepsilon})_{\lambda>0}) &= \Xi(\Omega\times\mathbb{R}^d, \mathbb{P}_V, \langle\cdot,\cdot\rangle_{n,\varepsilon}, \langle\cdot,\cdot\rangle_{n,\varepsilon}),\\ (\mathbb{H}_{n,\varepsilon}, L^{n,\varepsilon}, \mathrm{Dom}(L^{n,\varepsilon}), (G_\lambda^{L,n,\varepsilon})_{\lambda>0}) &= \Xi(\Omega\times\mathbb{R}^d, \mathbb{P}_V, \langle\cdot,\cdot\rangle_{n,\varepsilon}, B_{n,\varepsilon}).\end{aligned}$$

and we denote by $(L^{n,\varepsilon})^*$ the adjoint operator of $L^{n,\varepsilon}$ in $L^2(\Omega\times\mathbb{R}^d;\mathbb{P}_V)$.

Let us now consider a family $(\boldsymbol{\Psi}_\varepsilon)_\varepsilon$ of functions in $L^2(\Omega\times\mathbb{R}^d;\mathbb{P}_V)$ satisfying (37) for some family $(C_\varepsilon)_{\varepsilon>0}$ such that $\varepsilon C_\varepsilon \to 0$ as $\varepsilon \to 0$. Fix $n \in \bar{\mathbb{N}}^*$. Define $\boldsymbol{\varphi}_{n,\varepsilon} \equiv G_{\varepsilon^2}^{S,n,\varepsilon}(\boldsymbol{\Psi}_\varepsilon)$, which satisfies $\varepsilon^2\mathbb{M}_V[\boldsymbol{\varphi}_{n,\varepsilon}\boldsymbol{\psi}] + \langle\boldsymbol{\varphi}_{n,\varepsilon}, \boldsymbol{\psi}\rangle_{n,\varepsilon} = \mathbb{M}_V[\boldsymbol{\Psi}_\varepsilon\boldsymbol{\psi}]$ for any $\boldsymbol{\psi} \in \mathbb{H}_{n,\varepsilon}$. Choosing $\boldsymbol{\psi} = \boldsymbol{\varphi}_{n,\varepsilon}$, using (37) and the standard estimate $ab \le a^2/2 + b^2/2$ leads to

$$\begin{aligned}\varepsilon^2\mathbb{M}_V[|\boldsymbol{\varphi}_{n,\varepsilon}|^2] + \|\boldsymbol{\varphi}_{n,\varepsilon}\|_{n,\varepsilon}^2 = \mathbb{M}_V[\boldsymbol{\Psi}_\varepsilon\boldsymbol{\varphi}_{n,\varepsilon}] \le C_\varepsilon\sqrt{2}\|\boldsymbol{\varphi}_{n,\varepsilon}\|_{0,\varepsilon} &\le \sqrt{2}C_\varepsilon\|\boldsymbol{\varphi}_{n,\varepsilon}\|_{n,\varepsilon}\\ &\le C_\varepsilon^2 + \|\boldsymbol{\varphi}_{n,\varepsilon}\|_{n,\varepsilon}^2/2\end{aligned}$$

in such a way that

$$\varepsilon^2\mathbb{M}_V[|\boldsymbol{\varphi}_{n,\varepsilon}|^2] + \|\boldsymbol{\varphi}_{n,\varepsilon}\|_{n,\varepsilon}^2/2 \le C_\varepsilon^2. \tag{43}$$

Once again, to apply the Itô formula, we use vanishing viscosity methods in order to get round the lack of regularity of $\boldsymbol{\varphi}_{n,\varepsilon}$ because of the degeneracy of $\boldsymbol{a}$. In the non-degenerate framework ($n \ge 1$), from [13, Proof of Lemma 6.3], standard convolution technics provide us with a $\mathbb{H}_{n,\varepsilon}$-sequence $(\boldsymbol{\varphi}_{n,\varepsilon}^m)_{m\in\mathbb{N}}$ of smooth functions, namely that for each fixed $\omega \in \Omega$ the

function $x \mapsto \boldsymbol{\varphi}^m_{n,\varepsilon}(\tau_{x/\varepsilon}\omega, x)$ is a $C^\infty(\mathbb{R}^d)$-function, such that $\mathbb{M}_V[|\boldsymbol{\varphi}^m_{n,\varepsilon} - \boldsymbol{\varphi}_{n,\varepsilon}|^2 + |S^{n,\varepsilon}\boldsymbol{\varphi}^m_{n,\varepsilon} - S^{n,\varepsilon}\boldsymbol{\varphi}_{n,\varepsilon}|^2] + \|\boldsymbol{\varphi}^m_{n,\varepsilon} - \boldsymbol{\varphi}_{n,\varepsilon}\|^2_{n,\varepsilon} \to 0$ as m goes to ∞.

We are now going to use a time reversal argument. Let us consider the process (introduced in Section 7)

$$X^{n,\varepsilon}_t = x + \int_0^t (\varepsilon^{-1}b + c - n^{-1}\partial_y V)(\omega, \overline{X}^{n,\varepsilon}_r, X^{n,\varepsilon}_r)\,dr + \int_0^t \sigma(\omega, \overline{X}^{n,\varepsilon}_r, X^{n,\varepsilon}_r)\,dB_r + (n/2)^{-1/2}\tilde{B}_t,$$

where $\overline{X}^{n,\varepsilon}_r = X^{n,\varepsilon}_r/\varepsilon$. As explained in Section 7, its generator coincides on $C^2(\mathbb{R}^d)$ with

$$\mathcal{L}^{n,\varepsilon} = \frac{e^{2V(x)}}{2}\sum_{i,j}\partial_{x_i}\big(e^{-2V(x)}(a + H + n^{-1}\mathrm{Id})_{ij}(\omega, x/\varepsilon, x)\partial_{x_j}\,\cdot\,\big)$$

and admits $e^{-2V(x)}\,dx$ as invariant measure. Furthermore, for a fixed $T > 0$, the generator of the time reversed process $t \mapsto X^{n,\varepsilon}_{T-t}$ with initial law $e^{-2V(x)}\,dx$ coincides with the adjoint of $\mathcal{L}^{n,\varepsilon}$ in $L^2(\mathbb{R}^d; e^{-2V(x)}\,dx)$. For each $\varphi \in C^2(\mathbb{R}^d)$, it exactly matches

$$(\mathcal{L}^{n,\varepsilon})^*\varphi = \frac{e^{2V(x)}}{2}\sum_{i,j}\partial_{x_i}\big(e^{-2V(x)}(a - H + n^{-1}\mathrm{Id})_{ij}(\omega, x/\varepsilon, x)\partial_{x_j}\varphi\big)$$

As a consequence, observe that, for any $0 \le s \le t \le T$,

$$\begin{aligned}\varphi^m_{n,\varepsilon}(\overline{X}^{n,\varepsilon}_t, X^{n,\varepsilon}_t) =& \varphi^m_{n,\varepsilon}(\overline{X}^{n,\varepsilon}_s, X^{n,\varepsilon}_s) + \int_s^t [\mathcal{L}^{n,\varepsilon}(\varphi^m_{n,\varepsilon}(\cdot/\varepsilon, \cdot))](\overline{X}^{n,\varepsilon}_r, X^{n,\varepsilon}_r)\,dr\\ &+ (\overrightarrow{\mathcal{M}}^{m,n,\varepsilon}_t - \overrightarrow{\mathcal{M}}^{m,n,\varepsilon}_s),\end{aligned}$$

where $\overrightarrow{\mathcal{M}}^{m,n,\varepsilon}$ is a martingale with respect to the forward filtration $(\mathcal{F}^{n,\varepsilon}_t)_{0\le t\le T}$ and $\mathcal{F}^{n,\varepsilon}_t$ is the σ-algebra on $\mathbb{R}^d$ generated by $\{X^{n,\varepsilon}_r; 0 \le r \le t\}$. In the same way,

$$\begin{aligned}\varphi^m_{n,\varepsilon}(\overline{X}^{n,\varepsilon}_s, X^{n,\varepsilon}_s) =& \varphi^m_{n,\varepsilon}(\overline{X}^{n,\varepsilon}_t, X^{n,\varepsilon}_t) + \int_s^t [(\mathcal{L}^{n,\varepsilon})^*(\varphi^m_{n,\varepsilon}(\cdot/\varepsilon, \cdot))](\overline{X}^{n,\varepsilon}_r, X^{n,\varepsilon}_r)\,dr\\ &+ (\overleftarrow{\mathcal{M}}^{m,n,\varepsilon}_t - \overleftarrow{\mathcal{M}}^{m,n,\varepsilon}_s),\end{aligned}$$

where $\overleftarrow{\mathcal{M}}^{m,n,\varepsilon}$ is a martingale with respect to the backward filtration $(\mathcal{G}^{n,\varepsilon}_t)_{0\le t\le T}$ and $\mathcal{G}^{\varepsilon}_s$ is the σ-algebra on $\mathbb{R}^d$ generated by $\{X^{n,\varepsilon}_r; t \le r \le T\}$. Add these two expressions:

$$-2\varepsilon^{-2}\int_s^t S^{n,\varepsilon}\varphi^m_{n,\varepsilon}(\overline{X}^{n,\varepsilon}_r, X^{n,\varepsilon}_r)\,dr = (\overrightarrow{\mathcal{M}}^{m,n,\varepsilon}_t - \overrightarrow{\mathcal{M}}^{m,n,\varepsilon}_s) + (\overleftarrow{\mathcal{M}}^{m,n,\varepsilon}_t - \overleftarrow{\mathcal{M}}^{m,n,\varepsilon}_s).$$

We further mention that the quadratic variations of both martingales exactly match

$$\varepsilon^{-2}\int_s^t [(D + \varepsilon\partial_y)\varphi^m_{n,\varepsilon}]^* a[(D + \varepsilon\partial_y)\varphi^m_{n,\varepsilon}]^*(\overline{X}^{n,\varepsilon}_r, X^{n,\varepsilon}_r)\,dr,$$

in such a way that the Doob inequality yields

$$\bar{\mathbb{E}}^{\varepsilon}\Big[\sup_{0\leq s\leq t}|\int_0^s S^{n,\varepsilon}\varphi^m_{n,\varepsilon}(\overline{X}^{n,\varepsilon}_r, X^{n,\varepsilon}_r)\,dr|^2\Big]\leq 16T\varepsilon^2\|\boldsymbol{\varphi}^m_{n,\varepsilon}\|^2_{n,\varepsilon}.$$

Letting m go to ∞, reminding that $\varepsilon^2\boldsymbol{\varphi}_{n,\varepsilon}-S^{n,\varepsilon}\boldsymbol{\varphi}_{n,\varepsilon}=\boldsymbol{\Psi}_\varepsilon$ and using (43) leads to

$$\bar{\mathbb{E}}^{\varepsilon}\big[\sup_{0\leq s\leq t}|\int_0^s \Psi_\varepsilon(\overline{X}^{n,\varepsilon}_r, X^{n,\varepsilon}_r)\,dr|^2\big]\leq 32T\varepsilon^2\|\boldsymbol{\varphi}_{n,\varepsilon}\|^2_{n,\varepsilon}+2T\varepsilon^4\mathbb{M}_V[|\boldsymbol{\varphi}_{n,\varepsilon}|^2]\leq 68T\varepsilon^2C^2_\varepsilon.$$

We then complete the proof in letting n go to ∞ and in using the fact that $X^{n,\varepsilon}$ converges in $C([0,T];\mathbb{R}^d)$ towards X^ε as n goes to ∞. □

9 Proof of Theorem 3.1 and Proposition 3.2

Proof of Theorem 3.1. Section 10 below is devoted to proving the tightness of the family of processes $(X^\varepsilon)_\varepsilon$ in $C([0,T];\mathbb{R}^d)$. It remains to prove that there is a unique possible weak limit for all converging subsequences.

From now on, the corrector $\boldsymbol{u}^i_\lambda$ ($\lambda>0$ and $1\leq i\leq d$) stands for the solution of (20). Applying the Ito formula (Theorem 7.2) to the correctors leads to

$$\begin{aligned}dX^\varepsilon_t=&-\varepsilon du_{\varepsilon^2}(\overline{X}^\varepsilon_t,X^\varepsilon_t)+\varepsilon(\partial_y u_{\varepsilon^2})^*\sigma(\overline{X}^\varepsilon_t,X^\varepsilon_t)\,dB_t\\&+[\varepsilon u_{\varepsilon^2}+\varepsilon c\cdot\partial_y u_{\varepsilon^2}+(\varepsilon/2)\mathrm{trace}(a\partial^2_{yy}u_{\varepsilon^2})](\overline{X}^\varepsilon_t,X^\varepsilon_t)\,dt\\&+[b\partial_y u_{\varepsilon^2}+c\cdot(I+Du_{\varepsilon^2})+\mathrm{trace}(aD\partial_y u_{\varepsilon^2})](\overline{X}^\varepsilon_t,X^\varepsilon_t)\,dt\\&+[\sigma+Du_{\varepsilon^2}\sigma](\overline{X}^\varepsilon_t,X^\varepsilon_t)\,dB_t\\\equiv&\, d\Theta^{1,\varepsilon}_t+d\Theta^{2,\varepsilon}_t+d\Theta^{3,\varepsilon}_t+d\Theta^{4,\varepsilon}_t\end{aligned}$$

Concerning the first term, we have $\bar{\mathbb{E}}^\varepsilon\big[|\Theta^{1,\varepsilon}_t|^2\big]\leq(1+T)\varepsilon^2\mathbb{M}_V\big[|\boldsymbol{u}_{\varepsilon^2}|^2+M^2|\partial_y\boldsymbol{u}_{\varepsilon^2}|^2\big]$ for $0\leq t\leq T$. This latter quantity converges to 0 as ε goes to 0 from Proposition 6.3. The same job can be carried out for $\Theta^{2,\varepsilon}$ and the same conclusion holds.

The main difficulty actually lies in the term $\Theta^{3,\varepsilon}$, especially in the part corresponding to $\boldsymbol{b}\partial_y\boldsymbol{u}_{\varepsilon^2}$. Concerning the remaining part $\boldsymbol{c}\cdot(I+D\boldsymbol{u}_{\varepsilon^2})+\mathrm{trace}(\boldsymbol{a}D\partial_y\boldsymbol{u}_{\varepsilon^2})$, it is readily seen (see Proposition 6.3) that it converges in $L^2(\Omega\times\mathbb{R}^d;\mathbb{P}_V)$ and thus Theorem 8.1 can be applied. As a consequence, we have

$$\bar{\mathbb{E}}^{\varepsilon}\big[\sup_{0\leq t\leq T}\big|\int_0^t[c\cdot(I+Du_{\varepsilon^2})+\mathrm{trace}(aD\partial_y u_{\varepsilon^2})](\overline{X}^\varepsilon_r,X^\varepsilon_r)\,dr-\int_0^t\bar{\Phi}(X^\varepsilon_r)\,dr\big|^2\big]\to 0 \text{ as } \varepsilon\to 0,$$

where $\bar{\Phi}(y)=\lim_{\lambda\to 0}\mathbb{M}[\boldsymbol{c}\cdot(I+D\boldsymbol{u}_\lambda)+\mathrm{trace}(\boldsymbol{a}D\partial_y\boldsymbol{u}_\lambda)(\cdot,y)]$. It remains to treat the term $(\boldsymbol{b}\partial_y\boldsymbol{u}_{\varepsilon^2})_\varepsilon$. Note that the L^2-norm of $\boldsymbol{b}\partial_y\boldsymbol{u}_{\varepsilon^2}$ need not be convergent. That is why we have in

mind to use Theorem 8.2. Up to introducing new correctors, we will prove that $\boldsymbol{b}\partial_y \boldsymbol{u}_{\varepsilon^2}$ can be divided into two parts, satisfying respectively Theorems 8.1 and 8.2. To understand how this decomposition occurs, let us consider a test function $\boldsymbol{\varphi} \in \mathcal{C}_\Pi$. Then two successive integrations by parts yield, for $1 \le i, j \le d$, (we use the convention of summation over repeated indices)

$$\begin{aligned}
\mathbb{M}_V\big[\boldsymbol{b}_j \partial_{y_j} \boldsymbol{u}^i_{\varepsilon^2} \boldsymbol{\varphi}\big] &= (1/2)\mathbb{M}_V\big[D_p(\boldsymbol{a}+\boldsymbol{H})_{pj}\partial_{y_j}\boldsymbol{u}^i_{\varepsilon^2}\boldsymbol{\varphi}\big] \\
&= -(1/2)\mathbb{M}_V\big[(\boldsymbol{a}+\boldsymbol{H})_{pj}(D_p\partial_{y_j}\boldsymbol{u}^i_{\varepsilon^2}\boldsymbol{\varphi} + \partial_{y_j}\boldsymbol{u}^i_{\varepsilon^2} D_p\boldsymbol{\varphi})\big] \\
&= -(1/2)\mathbb{M}_V\big[(\boldsymbol{a}+\boldsymbol{H})_{pj}\big(D_p\partial_{y_j}\boldsymbol{u}^i_{\varepsilon^2}\boldsymbol{\varphi} + \partial_{y_j}\boldsymbol{u}^i_{\varepsilon^2}(D_p+\varepsilon\partial_{y_p})\boldsymbol{\varphi}\big)\big] \\
&\quad +(\varepsilon/2)\mathbb{M}_V\big[(\boldsymbol{a}+\boldsymbol{H})_{pj}\partial_{y_j}\boldsymbol{u}^i_{\varepsilon^2}\partial_{y_p}\boldsymbol{\varphi}\big] \\
&= -(1/2)\mathbb{M}_V\big[(\boldsymbol{a}+\boldsymbol{H})_{pj}D_p\partial_{y_j}\boldsymbol{u}^i_{\varepsilon^2}\boldsymbol{\varphi}\big] - (1/2)\mathbb{M}_V\big[\partial_{y_j}\boldsymbol{u}^i_{\varepsilon^2}(D_p+\varepsilon\partial_{y_p})\boldsymbol{\varphi}\big] \\
&\quad -(\varepsilon/2)\mathbb{M}_V\big[\partial_{y_p}(\boldsymbol{a}+\boldsymbol{H})_{pj}\partial_{y_j}\boldsymbol{u}^i_{\varepsilon^2}\boldsymbol{\varphi} + (\boldsymbol{a}+\boldsymbol{H})_{pj}\partial^2_{y_j y_p}\boldsymbol{u}^i_{\varepsilon^2}\boldsymbol{\varphi}\big] \\
&\quad +\varepsilon\mathbb{M}_V\big[(\boldsymbol{a}+\boldsymbol{H})_{pj}\partial_{y_j}\boldsymbol{u}^i_{\varepsilon^2}\boldsymbol{\varphi}\partial_{y_p}V\big].
\end{aligned}$$

So, for $1 \le i \le d$, define the correcting part $\mathrm{Corr}^i_\varepsilon(\omega, y) = (\varepsilon/2)\partial_{y_p}(\boldsymbol{a}+\boldsymbol{H})_{pj}\partial_{y_j}\boldsymbol{u}^i_{\varepsilon^2} + (\varepsilon/2)(\boldsymbol{a}+\boldsymbol{H})_{pj}\partial^2_{y_j y_p}\boldsymbol{u}^i_{\varepsilon^2} - \varepsilon(\boldsymbol{a}+\boldsymbol{H})_{pj}\partial_{y_j}\boldsymbol{u}^i_{\varepsilon^2}\partial_{y_p}V$, the L^2-converging part $\mathrm{Conv}^i_\varepsilon(\omega, y) = -(1/2)(\boldsymbol{a}+\boldsymbol{H})_{pj}D_p\partial_{y_j}\boldsymbol{u}^i_{\varepsilon^2}$ and L^2-diverging part $\mathrm{Div}^i_\varepsilon(\omega, y) = [\boldsymbol{b}_j\partial_{y_j}\boldsymbol{u}^i_{\varepsilon^2} + \mathrm{Corr}^i_\varepsilon - \mathrm{Conv}^i_\varepsilon](\omega, y)$. From the previous calculation, $\mathrm{Div}^i_\varepsilon$ satisfies the "Poincaré inequality" (37), namely that for any function $\boldsymbol{\varphi}$ in $\mathcal{C}_\Pi$, $\mathbb{M}_V\big[\boldsymbol{\Psi}_\varepsilon\boldsymbol{\varphi}\big] \le \big(M_V[|\partial_y \boldsymbol{u}^i_{\varepsilon^2}|^2]\big)^{1/2}\big(M_V[|(D+\varepsilon\partial_y)\boldsymbol{\varphi}|^2]\big)^{1/2}$. Moreover, Proposition 6.3 ensures that $\varepsilon\big(M_V[|\partial_y\boldsymbol{u}^i_{\varepsilon^2}|^2]\big)^{1/2} \to$ as ε goes to 0. Consequently, (38) holds for $\mathrm{Div}^i_\varepsilon$. Thanks to Proposition 6.3, the family $(\mathrm{Corr}^i_\varepsilon)_\varepsilon$ converges in $L^2(\Omega\times\mathbb{R}^d;\mathbb{P}_V)$ towards 0. As a consequence, $\bar{\mathbb{E}}^\varepsilon\big[\big(\int_0^t \mathrm{Corr}^i_\varepsilon(\overline{X}^\varepsilon_r, X^\varepsilon_r)\,dr\big)^2\big]$ tends to 0 as ε goes to 0. Then, Theorem 8.1 ensures that $\bar{\mathbb{E}}^\varepsilon\big[\sup_{0\le t\le T}|\int_0^t \mathrm{Conv}^i_\varepsilon(\overline{X}^\varepsilon, X^\varepsilon_r)\,dr - \int_0^t \overline{\Gamma}(X^\varepsilon_r)\,dr|^2\big] \to 0$ as $\varepsilon \to 0$, where $\overline{\boldsymbol{\Gamma}}(y) \equiv \lim_{\lambda\to 0} -(1/2)\mathbb{M}\big[(\boldsymbol{a}+\boldsymbol{H})_{pj}D_p\partial_{y_j}\boldsymbol{u}^i_\lambda(\cdot, y)\big]$. To sum up, this proves that

$$\bar{\mathbb{E}}^\varepsilon\big[\sup_{0\le t\le T}|\int_0^t b\cdot\partial_y u^i_{\varepsilon^2}(\overline{X}^\varepsilon, X^\varepsilon_r)\,dr - \int_0^t \overline{\Gamma}(X^\varepsilon_r)\,dr|^2\big] \to 0 \tag{44}$$

as ε tends to 0.

Concerning the martingale part $\Theta^{4,\varepsilon}$, it suffices to apply Theorem 8.1 to the quadratic variations.

Hence each possible limit point X in $C[0,t];\mathbb{R}^d)$ of the process X^ε must solve the martingale problem $X_t = x + \int_0^t \overline{B}(X_r)\,dr + \int_0^t \overline{A}^{1/2}(X_r)\,dB_r$, where the entries of $\bar{B}$ are given by

$$\begin{aligned}
\overline{B}_i &= \lim_{\lambda\to 0}\mathbb{M}\big[-(1/2)(\boldsymbol{a}+\boldsymbol{H})_{pj}D_p\partial_{y_j}\boldsymbol{u}^i_\lambda + \boldsymbol{c}_j(\delta_{ij}+D_j\boldsymbol{u}^i_\lambda) + \boldsymbol{a}_{pj}D_j\partial_{y_p}\boldsymbol{u}^i_\lambda\big] \\
&= \lim_{\lambda\to 0}\mathbb{M}\big[(1/2)(\boldsymbol{a}+\boldsymbol{H})_{pj}D_p\partial_{y_j}\boldsymbol{u}^i_\lambda + \boldsymbol{c}_j(\delta_{ij}+D_j\boldsymbol{u}^i_\lambda)\big] \\
&= \frac{e^{2V}}{2}\partial_{y_j}\big(e^{-2V}\lim_{\lambda\to 0}\mathbb{M}\big[(\boldsymbol{a}+\boldsymbol{H})_{pj}(\delta_{ij}+D_p\boldsymbol{u}^i_\lambda)\big]\big)
\end{aligned}$$

Thanks to Proposition 3.2, it is readily seen that the coefficients $\overline{B}$ and $\overline{A}^{1/2}$ are two times continuously differentiable with bounded derivatives up to order two. In particular, they are Lipschitzian and there exists a unique solution to the corresponding martingale problem. $\square$

Proof of Proposition 3.2. The strategy consists in introducing the homogenized diffusion coefficient associated to the operator $\widetilde{\boldsymbol{S}}$ and in comparing it with $\bar{A}(y)$. So we define the $d \times d$ nonnegative symmetric matrix $\widetilde{A}$ as the unique symmetric matrix satisfying (this is the classical variational formula for the homogenized coefficient associated to $\widetilde{\boldsymbol{S}}$, see [10] for further details)

$$\forall x \in \mathbb{R}^d, \quad \langle x, \widetilde{A}x\rangle = \inf_{\boldsymbol{\varphi}\in\mathcal{C}} \mathbb{M}\big[|\widetilde{\boldsymbol{\sigma}}^*(D\boldsymbol{\varphi}+x)|^2\big]. \tag{45}$$

Due to Assumption 2.4, we have for each function $\boldsymbol{\varphi} \in \mathcal{C}$,

$$M^{-1}\langle x, \widetilde{A}x\rangle \le M^{-1}\mathbb{M}\big[|\widetilde{\boldsymbol{\sigma}}^*(D\boldsymbol{\varphi}+x)|^2\big] \le \mathbb{M}\big[|\boldsymbol{\sigma}^*(\cdot,y)(D\boldsymbol{\varphi}+x)|^2\big].$$

Since $\mathcal{C}$ is dense in $\mathbb{H}_1$, we can choose $\boldsymbol{\varphi} = \boldsymbol{u}_\lambda(\cdot,y)\cdot x$ and then pass to the limit as λ tends to 0. We obtain $M^{-1}\langle x, \widetilde{A}x\rangle \le \langle x, \overline{A}(y)x\rangle$.

Now we turn to the auxiliary problems (subsection 6). Denoting by $\mathbb{L}$ the closure of $\{\tilde{\boldsymbol{\sigma}}^*\boldsymbol{\zeta},\ \boldsymbol{\zeta} \in L^2(\Omega;\mathbb{R}^d)\}$, we can extend $\boldsymbol{T}^y$ to the whole $\mathbb{L}$ as follows

$$\forall \boldsymbol{\zeta},\boldsymbol{\theta} \in L^2(\Omega,\mathbb{R}^d),\ \boldsymbol{T}^y(\tilde{\boldsymbol{\sigma}}^*\boldsymbol{\zeta},\tilde{\boldsymbol{\sigma}}^*\boldsymbol{\theta}) = (1/2)\big([\boldsymbol{a}+\boldsymbol{H}](\cdot,y)\boldsymbol{\zeta},\boldsymbol{\theta}\big)_2. \tag{46}$$

The underlying quadratic form is still denoted by $\boldsymbol{T}^y(\cdot)$. Furthermore, from Assumption 2.4, for some positive constant C only depending on M, we have

$$\boldsymbol{T}^y(\tilde{\boldsymbol{\sigma}}^*\boldsymbol{\zeta},\tilde{\boldsymbol{\sigma}}^*\boldsymbol{\theta}) \le C\boldsymbol{T}^y(\tilde{\boldsymbol{\sigma}}^*\boldsymbol{\zeta})^{1/2}\boldsymbol{T}^y(\tilde{\boldsymbol{\sigma}}^*\boldsymbol{\theta})^{1/2}. \tag{47}$$

Equation (24) then reads, for any function $\boldsymbol{\varphi} \in \mathcal{C}$,

$$\forall x \in \mathbb{R}^d,\ \boldsymbol{T}^y(\widetilde{\boldsymbol{\xi}}(\cdot,y)x,\tilde{\boldsymbol{\sigma}}^*\boldsymbol{D}\boldsymbol{\varphi}) = -(1/2)\big([\boldsymbol{a}+\boldsymbol{H}](\cdot,y)x,\boldsymbol{D}\varphi\big)_2 = -\boldsymbol{T}^y(\tilde{\boldsymbol{\sigma}}^*x,\tilde{\boldsymbol{\sigma}}^*\boldsymbol{D}\boldsymbol{\varphi}). \tag{48}$$

From (10a), (46) and (48), we have for any function $\boldsymbol{\varphi} \in \mathcal{C}$

$$\begin{aligned}\langle x, \overline{A}(y)x\rangle &= 2\lim_{\lambda\to 0}\boldsymbol{T}^y\big(\widetilde{\boldsymbol{\sigma}}^*x + \nabla^{\widetilde{\sigma}}\boldsymbol{u}_\lambda(\cdot,y)x\big) = 2\boldsymbol{T}^y(\widetilde{\boldsymbol{\sigma}}^*x + \widetilde{\boldsymbol{\xi}}(\cdot,y)x)\\ &= 2\boldsymbol{T}^y(\widetilde{\boldsymbol{\sigma}}^*x + \widetilde{\boldsymbol{\xi}}(\cdot,y)x, \widetilde{\boldsymbol{\sigma}}^*x + \widetilde{\boldsymbol{\sigma}}^*\boldsymbol{D}\boldsymbol{\varphi})\\ &\le 2C\boldsymbol{T}^y(\widetilde{\boldsymbol{\sigma}}^*x + \widetilde{\boldsymbol{\xi}}(\cdot,y)x)^{1/2}\boldsymbol{T}^y(\widetilde{\boldsymbol{\sigma}}^*x + \widetilde{\boldsymbol{\sigma}}^*\boldsymbol{D}\boldsymbol{\varphi})^{1/2}.\end{aligned}$$

Gathering this with the inequality $\boldsymbol{T}^y(\widetilde{\boldsymbol{\sigma}}^*x + \widetilde{\boldsymbol{\sigma}}^*\boldsymbol{D}\boldsymbol{\varphi}) \le M\mathbb{M}\big[|\widetilde{\boldsymbol{\sigma}}^*x + \widetilde{\boldsymbol{\sigma}}^*\boldsymbol{D}\boldsymbol{\varphi}|^2\big]$ and (45), we deduce $\langle x, \overline{A}(y)x\rangle \le 2C^2M\langle x, \widetilde{A}x\rangle$.

It just remains to prove that the drift term $\overline{B}$ is orthogonal to $K = \operatorname{Ker}\overline{A}(y)$. Due to (10c) and the fact that $K = \operatorname{Ker}\overline{A}(y)$ does not depend on $y \in \mathbb{R}^d$, it suffices to prove that $\operatorname{Ker}\overline{H}(y) \subset \operatorname{Ker}\overline{A}(y) = K$. But this is an easy consequence of (10a), (10b) and Assumption 2.4, especially $|\boldsymbol{H}(\omega,y)| \le M^2\boldsymbol{a}(\omega,y)$. $\square$

10 Tightness

We now turn to the tightness of the process X^ε, ie we want to prove that the family $(X^\varepsilon)_\varepsilon$ is tight in $C([0,T],\mathbb{R}^d)$ equipped with the uniform topology. That step of our result deeply differs from the uniform elliptic case [13]. Indeed, uniform ellipticity of the diffusion matrix provides strong transition density estimates of the process X^ε, the so-called Aronson estimates, from which the tightness of X^ε is then easily derived. Of course, in the degenerate framework, tightness of X^ε cannot be tackled this way. The method presented below is inspired from [15] and is based on the idea that the process X^ε is not too far from being reversible at a microscopic scale. The contributions of the macroscopic variations make a drift appear, unlike in [15].

Let us now go into details. As in Section 6, we can solve the following equation for $i = 1,\dots,d$ and $\lambda > 0$

$$\lambda \boldsymbol{w}^i_\lambda(.,y) - \boldsymbol{S}^y \boldsymbol{w}^i_\lambda(.,y) = \boldsymbol{b}_i(.,y) \tag{49}$$

and get the same properties as in Proposition 6.3, namely

Proposition 10.1. *For each fixed $y\in\mathbb{R}^d$ and $1\le i\le d$, the family $(\nabla^{\widetilde{\sigma}}\boldsymbol{w}^i_\lambda(.,y))_\lambda$ converges to a limit $\widetilde{\boldsymbol{\zeta}}_i(.,y)\in L^2(\Omega)^d$ as λ goes to 0. The same property holds for the derivatives, that is, the families $(\nabla^{\widetilde{\sigma}}\partial_{y_j}\boldsymbol{w}^i_\lambda)_\lambda$, $(\nabla^{\widetilde{\sigma}}\partial^2_{y_jy_k}\boldsymbol{w}^i_\lambda)_\lambda$ ($1\le i,j,k\le d$) respectively converge to $\partial_{y_j}\widetilde{\boldsymbol{\zeta}}_i(.,y)$, $\partial^2_{y_jy_jk}\widetilde{\boldsymbol{\zeta}}_i(.,y)$ in $L^2(\Omega)^d$. Furthermore, the function $\boldsymbol{w}^i_\lambda$ as well as its derivatives $\partial_{y_j}\boldsymbol{w}^i_\lambda$, $\partial^2_{y_jy_k}\boldsymbol{w}^i_\lambda$ satisfy* (6.3) *and estimates* (22) *and* (23)*, for some positive constant $C_{10.1}$ independent of $\lambda>0$ and $y\in\mathbb{R}^d$.*

As in the proof of Theorem 8.2, we want to use a time reversal argument. Once again, we are faced with the lack of smoothness of $\boldsymbol{w}_\lambda$ in order to apply the Itô formula. To overcome this difficulty, we proceed as in Section 7. Since the arguments are quite similar, we just outline the main ideas without further details. Let us consider, for $n\ge 1$, $\lambda>0$ and $1\le i\le d$, the solution $\boldsymbol{w}^{i,n}_\lambda$ of the following equation

$$\lambda \boldsymbol{w}^{i,n}_\lambda(.,y) - \boldsymbol{S}^y \boldsymbol{w}^{i,n}_\lambda(.,y) - n^{-1}\Delta \boldsymbol{w}^{i,n}_\lambda(.,y) = \boldsymbol{b}_i(.,y) \tag{50}$$

Introducing a sequence of regularizing sequence of mollifiers $(\varrho_m)_{m\in\mathbb{N}}\in C^\infty_c(\mathbb{R}^d\times\mathbb{R}^d)$ (smooth functions with compact support), we define

$$\boldsymbol{w}^{i,n}_{\lambda,m}(\omega,y) = \int_{\mathbb{R}^{2d}} \boldsymbol{w}^{i,n}_\lambda(\tau'_x\omega, y-y')\varrho_m(x',y')\,dx'\,dy',$$

which is a smooth function. Following the proof of Theorem 8.2, under the invariant measure

$e^{-2V(x)}\,dx$ of the process $X^{n,\varepsilon}$, we can write

$$
\begin{aligned}
(51)\qquad w^{i,n}_{\varepsilon^2,m}(\overline{X}^{n,\varepsilon}_t, X^{n,\varepsilon}_t) =& w^{i,n}_{\varepsilon^2,m}(\overline{X}^{n,\varepsilon}_s, X^{n,\varepsilon}_s) + \int_s^t [\mathcal{L}^{n,\varepsilon}(w^{i,n}_{\varepsilon^2,m}(\cdot/\varepsilon,\cdot))](\overline{X}^{n,\varepsilon}_r, X^{n,\varepsilon}_r)\,dr \\
&+ (\overrightarrow{\mathcal{M}}^{\varepsilon,n,m}_t - \overrightarrow{\mathcal{M}}^{\varepsilon,n,m}_s),\\
(52)\qquad w^{i,n}_{\varepsilon^2,m}(\overline{X}^{n,\varepsilon}_s, X^{n,\varepsilon}_s) =& w^{i,n}_{\varepsilon^2,m}(\overline{X}^{n,\varepsilon}_t, X^{n,\varepsilon}_t) + \int_s^t [(\mathcal{L}^{\varepsilon})^*(w^{i,n}_{\varepsilon^2,m}(\cdot/\varepsilon,\cdot))](\overline{X}^{n,\varepsilon}_r, X^{n,\varepsilon}_r)\,dr \\
&+ (\overleftarrow{\mathcal{M}}^{\varepsilon,n,m}_t - \overleftarrow{\mathcal{M}}^{\varepsilon,n,m}_s),
\end{aligned}
$$

where $\overrightarrow{\mathcal{M}}^{\varepsilon,n,m}$ and $\overleftarrow{\mathcal{M}}^{\varepsilon,n,m}$ are two martingales respectively with respect to the forward filtration $(\mathcal{F}^{n,\varepsilon}_s)_{0\le s\le T} \equiv \sigma\{X^{n,\varepsilon}_r; 0 \le r \le s\}$ and with respect to the backward filtration $(\mathcal{G}^{n,\varepsilon}_s)_{0\le s\le T} \equiv \sigma\{X^{n,\varepsilon}_r; s \le r \le T\}$. The quadratic variations of both martingales match

$$
\epsilon^{-2}\int_0^{\cdot} (D\boldsymbol{w}^{i,n}_{\varepsilon^2,m} + \varepsilon\partial_y \boldsymbol{w}^{i,n}_{\varepsilon^2,m})^*(a + n^{-1}\mathrm{Id})(D\boldsymbol{w}^{i,n}_{\varepsilon^2,m} + \varepsilon\partial_y \boldsymbol{w}^{i,n}_{\varepsilon^2,m})(\overline{X}^{n,\varepsilon}_r, X^{n,\varepsilon}_r)\,dr.
$$

Adding up (51) and (52), passing to the limit as $m\to\infty$ (as explained in [13, Lemma 5.3]) and $n\to\infty$ (as explained in Section 7) and using (49) leads to

$$
\begin{aligned}
(53)\qquad \varepsilon^{-1}\int_s^t b_i(\overline{X}^{\varepsilon}_r, X^{\varepsilon}_r)\,dr =& \int_s^t [\varepsilon w^i_{\varepsilon^2} + (1/2)\mathrm{trace}(aD\partial_y w^i_{\varepsilon^2})](\overline{X}^{\varepsilon}_r, X^{\varepsilon}_r)\,dr \\
&+ \int_s^t \frac{e^{2V}}{2}\big[div_y\big(e^{-2V}a[Dw^i_{\varepsilon^2} + \varepsilon\partial_y w^i_{\varepsilon^2}]\big)\big](\overline{X}^{\varepsilon}_r, X^{\varepsilon}_r)\,dr \\
&+ (1/2)\int_s^t \mathrm{Div}(a)\cdot\partial_y w^i_{\varepsilon^2}(\overline{X}^{\varepsilon}_r, X^{\varepsilon}_r)\,dr \\
&+ \varepsilon(\overrightarrow{\mathcal{M}}^{\varepsilon}_t - \overrightarrow{\mathcal{M}}^{\varepsilon}_s) + \varepsilon(\overleftarrow{\mathcal{M}}^{\varepsilon}_t - \overleftarrow{\mathcal{M}}^{\varepsilon}_s) \\
&\equiv E^{1,\varepsilon}_{s,t} + E^{2,\varepsilon}_{s,t} + T^{1,\varepsilon}_{s,t} + T^{2,\varepsilon}_{s,t},
\end{aligned}
$$

where $\varepsilon\overrightarrow{\mathcal{M}}^{\varepsilon}$ and $\varepsilon\overleftarrow{\mathcal{M}}^{\varepsilon}$ are two martingales, respectively with respect to the forward filtration $(\mathcal{F}^{\varepsilon}_s)_{0\le s\le T} \equiv \sigma\{X^{\varepsilon}_r; 0\le r\le s\}$ and with respect to the backward filtration $(\mathcal{G}^{\varepsilon}_s)_{0\le s\le T} \equiv \sigma\{X^{\varepsilon}_r; s\le r\le T\}$, with quadratic variations

$$
(54)\qquad \int_0^{\cdot} (D\boldsymbol{w}^i_{\varepsilon^2} + \varepsilon\partial_y \boldsymbol{w}^i_{\varepsilon^2})^* a (D\boldsymbol{w}^i_{\varepsilon^2} + \varepsilon\partial_y \boldsymbol{w}^i_{\varepsilon^2})(\overline{X}^{\varepsilon}_r, X^{\varepsilon}_r)\,dr.
$$

Theorem 8.1 establishes the following convergence

$$
\lim_{\varepsilon\to 0}\bar{\mathbb{E}}^{\varepsilon}\Big[\sup_{0\le t\le T}\Big|E^{1,\varepsilon}_{0,t} + E^{2,\varepsilon}_{0,t} - \int_0^t \bar{G}(X^{\varepsilon}_r)\,dr\Big|\Big] = 0,
$$

where

$$\bar{G}(y) = \mathbb{M}\big[(1/2)\mathrm{trace}(\boldsymbol{\sigma}\partial_y\boldsymbol{\xi}_i)(.,y) + (e^{2V}/2)\mathrm{div}_y\big(e^{-2V}\boldsymbol{\sigma}\boldsymbol{\xi}_i\big)(.,y)\big)\big].$$

From Proposition 10.1 and (22), $\bar{G}$ is bounded so that the tightness of the process $t \mapsto \int_0^t \bar{G}(X_r^\varepsilon)\,dr$ in $C([0,T],\mathbb{R})$ results from the Kolmogorov criterion. The tightness of $E^{1,\varepsilon} + E^{2,\varepsilon}$ follows.

Let us investigate now the term $T^{1,\varepsilon}_{s,t} = (1/2)\int_s^t \mathrm{Div}(a)\cdot\partial_y w^i_{\varepsilon^2}(\overline{X}^\varepsilon_r, X^\varepsilon_r)\,dr$. Note that it can not be treated with Theorem 8.1 because the L^2-norm of $\mathrm{Div}(\boldsymbol{a})\partial_y\boldsymbol{w}_{\varepsilon^2}$ need not be bounded. Inspired by the proof of Theorem 3.1 in Section 9, we define

$$\boldsymbol{\Psi}^i_\varepsilon \equiv \mathrm{Div}(\boldsymbol{a})\cdot\partial_y\boldsymbol{w}^i_{\varepsilon^2} + \mathrm{trace}(\boldsymbol{a}D\partial_y\boldsymbol{w}^i_{\varepsilon^2}) + \varepsilon div_y(\boldsymbol{a})\cdot\partial_y\boldsymbol{w}^i_{\varepsilon^2} + \varepsilon\mathrm{trace}(\boldsymbol{a}\partial^2_{yy}\boldsymbol{w}^i_{\varepsilon^2}) - 2\varepsilon\boldsymbol{a}_{pj}\partial_{y_j}\boldsymbol{u}^i_{\varepsilon^2}\partial_{y_p}V.$$

By making two successive integrations by parts as in Section 9, we establish for any $\boldsymbol{\varphi} \in \mathcal{C}\times C_0^\infty(\mathbb{R}^d)$:

$$\mathbb{M}_V[\boldsymbol{\Psi}^i_\varepsilon,\boldsymbol{\varphi}] = -\mathbb{M}_V[\boldsymbol{a}\partial_y\boldsymbol{w}^i_{\varepsilon^2}\cdot(D\boldsymbol{\varphi}+\varepsilon\partial_y\boldsymbol{\varphi})] \overset{\text{Prop. }10.1}{\leq} C_\varepsilon\mathbb{M}_V[|\boldsymbol{\sigma}^*(D\boldsymbol{\varphi}+\varepsilon\partial_y\boldsymbol{\varphi})|^2]^{1/2},$$

where the family $(\varepsilon C_\varepsilon)_\varepsilon$ converges to 0 as ε goes to 0. Theorem 8.2 then ensures that

$$\bar{\mathbb{E}}^\varepsilon\Big[\sup_{0\leq s\leq t}\Big(\int_s^t \Psi_\varepsilon(\overline{X}^\varepsilon_r, X^\varepsilon_r)\,dr\Big)^2\Big] \to 0$$

as ε goes to 0. Thanks to Theorem 8.1 and Proposition 10.1, we have

$$\bar{\mathbb{E}}^\varepsilon\Big[\sup_{0\leq s\leq t}\Big|\int_0^s \mathrm{trace}(\boldsymbol{a}D\partial_y\boldsymbol{w}_{\varepsilon^2})(\overline{X}^\varepsilon_r, X^\varepsilon_r)\,dr - \int_0^s \bar{\Phi}(X^\varepsilon_r)\,dr\Big|^2\Big] \to 0$$

as ε goes to 0, where $\bar{\Phi}(y) = \lim_{\varepsilon\to 0}\mathbb{M}[\mathrm{trace}(\boldsymbol{a}D\partial_y\boldsymbol{w}_{\varepsilon^2})(\cdot,y)]$. The Kolmogorov criterion and Proposition 10.1 ensure the tightness in $C([0,t];\mathbb{R})$ of the process $\int_0^\cdot \bar{\Phi}(X^\varepsilon_r)\,dr$. Moreover, from Proposition 10.1 and (30), the process $\int_0^\cdot \big[\varepsilon div_y(\boldsymbol{a})\cdot\partial_y\boldsymbol{w}^i_{\varepsilon^2} + \varepsilon\mathrm{trace}(\boldsymbol{a}\partial^2_{yy}\boldsymbol{w}^i_{\varepsilon^2}) - 2\varepsilon\boldsymbol{a}_{pj}\partial_{y_j}\boldsymbol{u}^i_{\varepsilon^2}\partial_{y_p}V\big](\overline{X}^\varepsilon_r, X^\varepsilon_r)\,dr$ converges in law in $C([0,T];\mathbb{R})$ to 0. This proves the tightness of $T^{1,\varepsilon}$ in $C([0,t];\mathbb{R}^d)$.

It just remains to treat the martingale term $T^{2,\varepsilon}$. According to Theorem 4.13 in [6], it suffices to establish the tightness of the brackets of these two martingales (see (54)). Their tightness results from Theorem 8.1, Proposition 10.1 and the Kolmogorov criterion again. The tightness of X^ε is now clear. □

References

[1] Benchérif-Madani A., Pardoux E., *Homogenization of a diffusion with locally periodic coefficients*, in Séminaire de Probabilités XXXVIII, 363–392, Lecture Notes in Math., 1857, Springer, Berlin, 2005.

[2] BENSOUSSAN A., LIONS J.L., PAPANICOLAOU G., *Asymptotic methods in periodic media*, Ed. North Holland, 1978.

[3] DELARUE F., RHODES R., *Stochastic homogenization of quasilinear PDEs with a spatial degeneracy*. Submitted

[4] FUKUSHIMA, M., *Dirichlet forms and Markov processes*, North-Holland Publishing Co., Amsterdam-New York; Kodansha, Ltd., Tokyo, 1980.

[5] HAIRER M., PARDOUX E., *Homogenization of periodic linear degenerate PDEs*, submitted.

[6] JACOD J., SHIRYAEV A.N., *Limit Theorems for Stochastic Processes*, Grundlehren der mathematischen Wissenschaft 288, Springer-Verlag, Berlin Heidelberg, 1987.

[7] JIKOV V.V., KOZLOV S.M., OLEINIK O.A., *Homogenization of differential operators and integral functionals*, Springer-Verlag, Berlin, 1994.

[8] KRYLOV N.V., *Controlled Diffusion Processes*, Springer-Verlag, New York, 1980.

[9] MA Z.M., RÖCKNER M., *Introduction to the Theory of (Non-Symmetric) Dirichlet Forms*, Universitext, Berlin Heidelberg, Springer-Verlag, 1992.

[10] OLLA S., *Homogenization of diffusion processes in Random Fields*, Cours de l'école doctorale, Ecole polytechnique, 1994.

[11] OLLA S., SIRI P., *Homogenization of a bond diffusion in a locally ergodic random environment*, Stochastic Processes and Applications, 2004, vol 109, 317-326.

[12] RHODES R., *On Homogenization of space time dependent random flows*, Stochastic Processes and their Applications, 2007, vol 117, 1561-1585.

[13] Rhodes R., *Diffusion in a Locally Stationary Random Environment*, to appear in Probability Theory and Related Fields (available online).

[14] STROOCK D., *Diffusion semi-groups corresponding to uniformly elliptic divergence form operators*, in Séminaires de Probabilités XXII, Lecture Notes in Mathematics 1321, 1988, p. 316-347. section B, 1999, 35, 121-141.

[15] WU L., *Forward-Backward martingale decomposition and compactness results for additive functionals of stationary ergodic Markov processes*, Ann. Inst. Henri Poincaré section B, 1999, 35, 121-141.